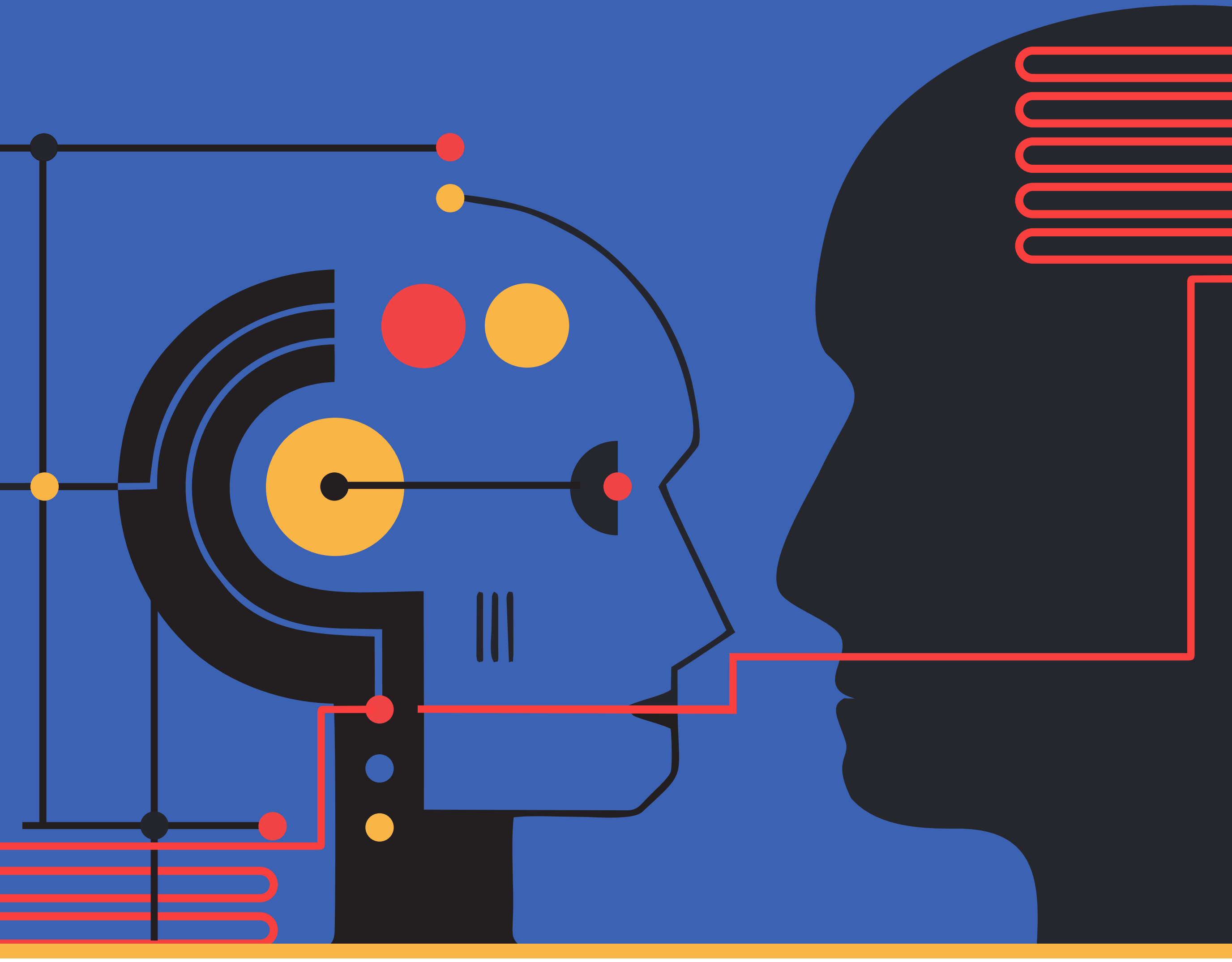
# ARTIFICIAL INTELLIGENCE

## and Postgraduate Supervision in Higher Education



ERRCDF
Rethinking Education 4 Inclusive Development

**Editors**
I. Kariyana & W. Sinkala

# ARTIFICIAL INTELLIGENCE

*and*

# Postgraduate Supervision in Higher Education

***Editors***
I. Kariyana & W. Sinkala

**ISBN:** 978-1-0492-7538-3 (e-book)
**ISBN:** 978-1-0492-7537-6 (print)
**Chapters' DOI:** https://doi.org/10.38140/obp4-2026c
**Book DOI:** https://doi.org/10.38140/obp4-2026
**Publication date:** March 2026
**Publisher:** ERRCD Forum
**Country:** South Africa

## COPYRIGHT AND LICENSING



## DISCLAIMER



## PEER REVIEW CONFIRMATION

The publisher, ERRCD Forum, and the editors affirm that every chapter included in this book underwent a rigorous double-blind peer-review process. This process ensured that each submission was independently reviewed by two experts in the field, with the identities of both the authors and reviewers kept confidential to maintain impartiality. Prior to the peer review, the editors assessed each chapter's quality and relevance to the book's scope. A total of 31 chapters were submitted for consideration in the book, of which 14 were ultimately accepted and published.

## PUBLISHER'S INFORMATION

**ERRCD Forum**
Harrismith, South Africa

This edited book is published via Open Books and Proceedings, an extension of the ERRCD Forum, committed to advancing knowledge through the publication of high-quality academic books, edited volumes, conference proceedings, and other scholarly monographs. Founded by a collective of experienced scholars, our mission is to ensure that cutting-edge research and evidence-based literature are accessible to all. ERRCD Forum is a registered private and independent publishing house that specialises in open-access publications. Our dedication to open access means that all our published works are freely available to readers, students, and researchers worldwide. We believe in the power of knowledge sharing and are passionate about making scholarly work accessible to a broader audience.

*OBP* is hosted, archived and maintained by the University of the Free State Library and Information Services and published by the ERRCD Forum. ERRCD Forum is based in South Africa and the United States and publishes open-access academic and peer-reviewed books and journals in education, rural and community studies, management and social sciences, among others.

# EDITORS' PROFILES

**Israel Kariyana**
Continuing Professional Teacher Development
Faculty of Education, Walter Sisulu University, South Africa
ORCID: https://orcid.org/0000-0002-5683-6814

Prof. Israel Kariyana is an Associate Professor of Education Management at Walter Sisulu University. He holds a Doctor of Education in Education Management and a Bachelor of Science (Honours) in Mathematics from Walter Sisulu University. He also holds a Master of Science in Strategic Management from Chinhoyi University of Technology and a Bachelor of Science in Education from the University of Zimbabwe. He is an interdisciplinary researcher with interests in educational development and quality enhancement, education management, educational administration, mathematics education, gamification in education, AI-enhanced postgraduate supervision, innovative digital pedagogies in higher education, and digital academic leadership. He leads various faculty and departmental committees, has project management skills, and is involved in national and international projects in various capacities. Prof. Kariyana has commendable publications and experience in supervision and community engagement. To date, he has successfully co-edited one book.

**Winter Sinkala**
Department of Mathematical Sciences and Computing
Faculty of Natural Sciences
Walter Sisulu University, Mthatha Campus, South Africa
ORCID: https://orcid.org/0000-0002-2263-7286

Prof. Sinkala is an Associate Professor of Applied Mathematics in the Department of Mathematical Sciences and Computing at Walter Sisulu University, South Africa. He holds a PhD in Mathematics from the University of KwaZulu-Natal and specialises in Lie Symmetry Analysis of Differential Equations. His research portfolio is structured around several key methodological themes within symmetry analysis, including innovative approaches to finding Lie point and non-classical symmetries, group classification of partial differential equations, construction of conservation laws, symmetry reductions via classical and double reduction methods leading to invariant solutions, integration of systems of first-order ordinary differential equations, and the construction of equivalence transformations to relate equivalent differential equations. He also has strong scholarly interests in Mathematics Education and computational fixed-point approaches. Professor Sinkala is actively involved in postgraduate supervision and academic leadership and has contributed to institutional quality assurance and research development initiatives.

## AUTHORS' PROFILES

**Maryam Chinasa Adewunmi**
Centre for Education Rights and Transformation (CERT)
Faculty of Education, University of Johannesburg, South Africa
ORCID: https://orcid.org/0009-0008-6735-8273

Maryam Chinasa Adewunmi is a curriculum scholar at the University of Johannesburg, affiliated with the Faculty of Education and the Centre for Education Rights and Transformation (CERT). She recently successfully defended her PhD in Curriculum Studies (degree to be conferred), where she examined the Africanisation of the International Relations Honours curriculum through African Indigenous Feminist Theory. She holds a BA in Political Leadership and Citizenship, an Honours degree in International Politics, a Postgraduate Certificate in Education, and both Honours and Master's degrees in Inclusive Education from the University of the Witwatersrand. Her interdisciplinary scholarship engages decoloniality, epistemic justice, African Indigenous knowledge systems, and socially responsive curriculum transformation in higher education.

**Dimitar Angelov**
Research Centre for Global Learning
Coventry University, United Kingdom
ORCID: https://orcid.org/0000-0003-3390-0163

Dr Dimitar Angelov is an Assistant Professor at the Research Centre for Global Learning at Coventry University, UK. A specialist in academic writing and writing for publication, he has led the conceptualisation and delivery of innovative researcher development interventions for early-, mid-career, and senior researchers in the UK and internationally. Dr Angelov's scholarly interests address the impact of artificial intelligence on higher education and knowledge production, as well as the ethics of transnational partnerships between institutions in the Global North and Global South. Dr Angelov holds a PhD in Literary Studies from the University of Warwick, UK.

**Krish Chetty**
Equitable Education and Economies Division
Human Sciences Research Council (HSRC), South Africa
ORCID: https://orcid.org/0000-0003-2731-3560

Dr Chetty is a Senior Research Manager at the Human Sciences Research Council (HSRC), South Africa. He holds a PhD in Information Systems from Nelson Mandela University and a Master's degree in Knowledge and Information Management from Stellenbosch University. His work centres on the social implementation of innovative technologies, particularly Artificial Intelligence, to address systemic challenges of inequality, poverty, and unemployment. With extensive experience in digital inclusion and knowledge-sharing within the Global South, his current research explores AI's role in good governance, education, and inclusive economic development.

**Tichaona Chikore**
Department of Mathematics and Applied Mathematics
Faculty of Science, University of Johannesburg, South Africa
ORCID: https://orcid.org/0000-0003-3209-5326

Dr Tichaona Chikore is an applied mathematician whose research centres on mathematical modelling, epidemiological systems, and data-driven analysis of complex socio-technical phenomena. He holds a PhD in Applied Mathematics and has contributed to modelling studies spanning infectious disease dynamics, behavioural transmission processes, and digital pathogen ecosystems. His work integrates dynamical systems theory, computation, and statistical inference to inform public health and technological risk decision-making. Dr Chikore is currently a researcher in the Department of Mathematics and Applied Mathematics at the University of Johannesburg, South Africa, where he collaborates on interdisciplinary modelling initiatives and postgraduate training in mathematical epidemiology and complex systems.

**Edmore Chinhamo**
Executive Business Centre,
Greenwich Business School, University of Greenwich, London.
ORCID: https://orcid.org/0000-0003-4684-7517

Dr Edmore Chinhamo is an academic with vast industry experience gained from over 25 years in private sector engagements. He holds a PhD in Operations Management from Nelson Mandela University, an MTech from Cape Peninsula University of Technology, a BCom Honours from Midlands State University, and a Diploma in Education from the University of Zimbabwe. He is also a Fellow of the Higher Education Academy (UK). His research focuses on Global Business Services, Operations Management, Global Value Chains, and Business Process Outsourcing. Dr Edmore Chinhamo is currently a lecturer at the University of Greenwich, London.

**Charles Bwalya Chisanga**
Department of Plant and Environmental Sciences
School of Natural Resources, The Copperbelt University, Zambia
ORCID: https://orcid.org/0000-0002-7388-5415

Dr Chisanga is a lecturer, researcher, and consultant in Soil Science, Climate and Crop Modelling, Wetland Mapping, Geographical Information Systems (GIS), Remote Sensing, and Climate Change Adaptation and Mitigation. He holds a PhD in Integrated Soil Fertility Management (ISFM) and specialises in soil fertility, climate modelling, crop modelling, hydrology, integrated soil fertility management, modelling ecosystem services, and geospatial applications, including remote sensing and GIS. His work integrates statistical and dynamical climate modelling to support climate-resilient agricultural systems and sustainable land management. He has contributed to major national and international climate policy processes, including Zambia's Biennial Transparency Report and National Communications to the UNFCCC, the National Adaptation Plan (NAP), and climate vulnerability assessments. Currently, he serves as a Lead Author for Working Group I of the IPCC Seventh Assessment Report and contributes to global climate research initiatives under the World Climate Research Programme, focusing on African monsoon systems, climate variability, and land–atmosphere interactions.

**Godwin Pedzisai Dzvapatsva**
Technology and Engineering Department
School of Business, Arts, Social Sciences and Technology
University of Suffolk, United Kingdom
ORCID:  https://orcid.org/0000-0001-5534-2647

Dr Dzvapatsva is a researcher with an interest in the use of ICT in education. He holds a PhD in Information Systems from the University of Cape Town. He is presently working as a lecturer in computing at the University of Suffolk, where he teaches undergraduate and postgraduate courses. Godwin’s research interests include ICT4D, mobile learning environments, and software engineering.

**Bramwell Kundishora Gavaza**
Department of Management Sciences
Faculty of Management & Public Administration Sciences
Walter Sisulu University, South Africa
ORCID: https://orcid.org/0000-0002-3684-5166

Dr Bramwell Kundishora Gavaza is a Senior Lecturer in the Department of Management Science at Walter Sisulu University. He holds a Doctor of Commerce in Business Management, a Master of Commerce in Business Management, a Master of Education, an Honours Degree in Business Management, a Postgraduate Diploma in Higher Education, a BTech in Business Administration, a BTech in Marketing, and a Diploma in Marketing. Driven, focused, and committed to excellence, Dr Gavaza has established a strong track record of delivering results both independently and collaboratively. His academic and research expertise spans Business Management, Entrepreneurship, and Marketing, with a particular emphasis on township economies, small business development, and educational technology. His work offers critical insights into the evolving landscape of township and rural development, contributing to debates on inclusive economic growth and sustainable entrepreneurial ecosystems. Through his teaching, research, and community engagement, he remains dedicated to advancing scholarship and fostering development-oriented innovation in South Africa and beyond.

**Zingiswa Jojo**
Secondary and Post School Department
Rhodes University, South Africa

ORCID: https://orcid.org/0000-0002-4949-1694

Prof Zingiswa Jojo is a renowned researcher and educator whose research interests spread across the fields of Ubuntu in the teaching and learning of geometric concepts, conceptual understanding of calculus concepts, instructional design in mathematics teaching, ethnomathematics, and indigenous knowledge in mathematics teaching and learning. She holds a Master's and PhD in Mathematics Education from the University of KwaZulu-Natal, South Africa.

**Israel Kariyana**
Continuing Professional Teacher Development
Faculty of Education, Walter Sisulu University, South Africa
ORCID: https://orcid.org/0000-0002-5683-6814

Prof. Israel Kariyana is an Associate Professor of Education Management at Walter Sisulu University. He holds a Doctor of Education in Education Management and a Bachelor of Science (Honours) in Mathematics from Walter Sisulu University. He also holds a Master of Science in Strategic Management from Chinhoyi University of Technology and a Bachelor of Science in Education from the University of Zimbabwe. He is an interdisciplinary researcher with interests in educational development and quality enhancement, education management, educational administration, mathematics education, gamification in education, AI-enhanced postgraduate supervision, innovative digital pedagogies in higher education, and digital academic leadership. He leads various faculty and departmental committees, has project management skills, and is involved in national and international projects in various capacities. Prof. Kariyana has commendable publications and experience in supervision and community engagement. To date, he has successfully co-edited one book.

**Nondumiso Masuku**
Human Sciences Research Council,
Equitable Economies and Education, Pretoria, South Africa.
ORCID: https://orcid.org/0009-0006-3933-1802

Nondumiso Masuku is a researcher in the Equitable Economies and Education unit at the Human Sciences Research Council, South Africa. She holds an Honours Bachelor of Commerce in Economics from the University of KwaZulu-Natal (2019). Her research interests focus on inclusive economic growth, with particular emphasis on labour market dynamics, youth employment, and skills development in the context of the Fourth Industrial Revolution.

**Courage Matobobo**
Department of Business and Application Development
Faculty of Engineering, Built Environment and Information Technology
Walter Sisulu University, South Africa
ORCID: https://orcid.org/0000-0002-7125-5989

Dr Courage Matobobo is a Senior Lecturer and researcher in Information Technology. He holds a PhD in Information Systems and has several years of experience in higher education, where he is actively involved in teaching, postgraduate supervision, curriculum development, and research. His academic work integrates both technical and socio-technical perspectives, with a strong focus on the impact of emerging technologies on society and education. His research interests include Information and Communication Technologies for Development (ICT4D), Technology in Education, Artificial Intelligence in Education, Information Systems design and implementation, and Social Media analytics. He is particularly interested in exploring how digital technologies can be leveraged to improve learning environments, enhance digital inclusion, and support sustainable development initiatives in developing contexts. Dr Matobobo has presented his research at numerous local and international conferences and has published in peer-reviewed journals, conference proceedings, and scholarly book chapters. He actively collaborates with interdisciplinary research teams and contributes to projects aimed at advancing digital transformation in education and community development. His scholarly work reflects a commitment to bridging the gap between technological innovation and real-world application.

**Ncediwe Mdlulwa**
Centre for Innovation in Learning and Teaching
Centre for Higher Education Development
University of Cape Town, South Africa
ORCID: https://orcid.org/0000-0002-9404-0841

Dr Ncediwe Mdlulwa is a lecturer based in the academic staff development cluster in the Centre for Innovation in Learning and Teaching (CILT) at the University of Cape Town (UCT). Dr Mdlulwa holds a PhD qualification within the Educational Psychology Department at the University of Johannesburg. Her current research explores the effectiveness of Artificial Intelligence (AI) in postgraduate supervision and its policy implications, as well as recommending AI solutions to strengthen youth employability interventions in South Africa. She endeavours to contribute towards augmenting teaching and learning practices in higher education by leveraging AI and innovative practices, drawing insights from the Global South.

**Luca Morini**
Centre for Global Learning
Coventry University, United Kingdom
ORCID: https://orcid.org/0000-0001-9391-0481

Dr Luca Morini is a researcher, educator and Postgraduate Programmes Lead at the Centre for Global Learning, Coventry University, United Kingdom. He holds a Master's in Psychology from the University of Bergamo and a PhD in Education and Communication, focusing on informal learning, online communities and co-creation, from the University of Milano Bicocca. His research focuses on research education and research engagement across different audiences, from doctoral candidates to young people to broader citizen communities, always with an emphasis on democracy and co-creation.

**Tirivanhu Muchuweni**
Department of Secondary and Post School Education
Rhodes University, South Africa
ORCID: https://orcid.org/0009-0005-7230-2803

Tirivanhu Muchuweni is a PhD candidate in Mathematics Education at Rhodes University, South Africa. His research focuses on gamification, artificial intelligence in education, and mathematics educational leadership, with an emphasis on student engagement, motivation, and learning outcomes. His work is broadly grounded in mathematics education across secondary and higher education contexts. He holds a BScEd Honours in Mathematics and an MSc in Mathematics Education from Bindura University of Science Education (BUSE). He has experience teaching mathematics and explores the use of digital and AI-supported tools to enhance teaching and learning.

**Obrain Tinashe Murire**
Department of Networking and Information Technology Support
Faculty of Engineering, Built Environment and Information Technology,
Walter Sisulu University, Buffalo City, South Africa
ORCID: https://orcid.org/0000-0003-3133-8744

Dr Obrain Tinashe Murire is a senior lecturer and a distinguished academic leader. Murire holds an honours degree, an MA, a PhD in Information Systems, and a postgraduate diploma (PGDip) in Higher Education, all from Fort Hare University. His research expertise spans several critical areas, including the 4IR, information security, emerging technologies in higher education, and youth employability. Murire has made significant contributions to both local and international academic discourse through numerous journal publications, peer-reviewed conference proceedings, and book chapters. His research is notable for its practical relevance, particularly in leveraging emerging technologies to improve educational outcomes and job prospects for young people. His work on the 4IR reflects his commitment to exploring how digital advancements are reshaping industries and employment landscapes, especially in developing countries. Beyond his research, Murire plays a key role in academic mentorship, supervising both master's and honours students and collaborating with colleagues on various research projects. His leadership as the research chairperson of his department has been instrumental in advancing the department's research agenda and fostering an environment of scholarly collaboration and innovation. Outside the university, he is actively involved in community initiatives, contributing to research-driven solutions that address local challenges. His community work has led to several journal publications, further solidifying his reputation as a scholar dedicated to bridging the gap between academia and real-world applications.

**Gardner Mwansa**
Department of Networking and Information Technology Support
Faculty of Engineering, Built Environment and Information Technology
Walter Sisulu University, East London, South Africa
ORCID: https://orcid.org/0000-0003-3827-9783

Dr Mwansa is a Senior Lecturer in the Department of Networking and Information Technology Support at Walter Sisulu University, South Africa. His research encompasses agile methodologies, cloud computing, artificial intelligence, the Internet of Things (IoT), cybersecurity, and educational technology. He concentrates on integrating agile practices within cloud environments, advancing secure IoT deployment, and applying AI-driven approaches to enhance cybersecurity and digital resilience in developing contexts. His research in educational technology investigates digital tools that improve learning outcomes and broaden access to e-learning. He is a Reviewing Editor within the Springer Nature reviewer community and serves on the editorial board of the POTRAZ Research Journal. He is a member of IEEE and the IEEE Computer Society, as well as a full member of the Information and Communication Technology Association of Zambia (ICTAZ).

**Matipa Ricky Ngandu**
Department of Networking and Information Technology Support
Faculty of Engineering, Built Environment and Information Technology
Walter Sisulu University, East London, South Africa
ORCID: https://orcid.org/0000-0003-4098-5070

Mr Ngandu is an academic with a research interest in cybersecurity, artificial intelligence (AI), and ICT4SD. He holds a Master of Philosophy (MPhil) degree in ICT from the University of Cape Town. He has taught undergraduate and postgraduate ICT modules and is committed to strengthening alignment between academia and industry to enhance graduate employability. He has published in recognised international journals and books and has presented at conferences in his research niche area. He is a professional member of the Institute of Information Technology Professionals South Africa (PMIITPSA), an organisation dedicated to upholding professional standards in the IT industry. This designation is SAQA-registered and IFIP IP3-accredited, reflecting his commitment to maintaining high levels of competence and integrity in the ICT sector.

**Prince Daughin Ngqabutho Ncube**
Department of Networking & IT Support
Faculty of Engineering, Built Environment & Information Technology
Walter Sisulu University, South Africa
ORCID: https://orcid.org/0000-0002-9262-3779

Prince Ncube is a lecturer and researcher working at the intersection of artificial intelligence, digital inclusion, and socio-technical systems in developing societies. His scholarship examines how AI-driven technologies reshape knowledge access, education systems, and data-informed governance, with particular attention to equity, responsible innovation, and technological adoption in low-resource environments. His research brings together computational methods, bibliometric analysis, and design-oriented approaches to develop AI-enabled tools and analytical frameworks addressing challenges in education, public safety, and digital transformation. Through this work, he contributes to emerging global conversations on responsible AI, algorithmic mediation of knowledge, and inclusive digital futures, with a strong focus on African contexts.

**Thembi Busisiwe Nkosi**
Provincial Department of Education: Quality Assurance (Quality Management and Support
Mpumalanga Department of Education
ORCID: https://orcid.org/0009-0003-7580-234X

Thembi Busisiwe Nkosi is an Educational Leadership Specialist in the Mpumalanga Department of Education's Directorate of Quality Assurance (Quality Management and Support Sub-Directorate). She supports primary and secondary schools province-wide to boost academic performance. Her qualifications include a Diploma in Education (Mgwenya College of Education), BA and Honours in Education (UNISA), and a Master's in Education Leadership and Management (University of Johannesburg). Her Master's thesis explored female principals' leadership in rural primary schools during COVID-19, spotlighting digital and infrastructural challenges alongside technology's transformative role. Thembi has just finalised her doctorate (awaiting graduation) on rural female principals' integration of Fourth Industrial Revolution (4IR) technologies in public primary schools. Findings highlight adaptive strategies like WhatsApp for community ties, e-learning for administration, and robotics for STEM, yielding the Rural 4IR Integration Framework (R4IRIF) for equitable 21st-century skills.

**Farai Nyabadza**
Department of Mathematics and Applied Mathematics
Faculty of Science
University of Johannesburg, South Africa
ORCID: https://orcid.org/0000-0003-3468-5581

Prof. Farai Nyabadza is a distinguished applied mathematician and academic leader specialising in mathematical epidemiology, dynamical systems, and mathematical modelling of socio-ecological and public health processes. He holds a PhD in Applied Mathematics and has held academic appointments in South Africa and internationally. His research focuses on infectious disease dynamics, behavioural–epidemiological interactions, digital and socio-environmental systems, and policy-relevant modelling for the Global South. Prof. Nyabadza currently serves as Professor and Head of the Department of Mathematics and Applied Mathematics at the University of Johannesburg, where he leads interdisciplinary modelling research and postgraduate supervision across epidemiology, digital risk systems, and sustainability applications.

**Vincent Raphael Nyirenda**
Department of Zoology and Aquatic Sciences
School of Natural Resources
The Copperbelt University, Zambia
ORCID: https://orcid.org/0000-0002-3506-3912

Prof. Nyirenda is a distinguished academic, researcher, and consultant specialising in conservation science, with academic qualifications including a Bachelor of Science in Natural Resource Management (University of Zambia, Zambia), a Master of Science in Environmental Systems Analysis and Management (University of Twente, the Netherlands), and a PhD in Nature Conservation (Tshwane University of Technology, South Africa). His research focuses on conservation science, human–nature interactions, social–ecological sustainability, sustainable tourism, and functional ecology, with a strong emphasis on interdisciplinary and evidence-based approaches to natural resource management. He has published widely in peer-reviewed journals and contributed to advancing conservation and sustainability science. His current scholarly interests also include innovative pedagogical techniques to enhance environmental and conservation education.

**Peter Babajide Oloba**
Department of Languages, Cultural Studies and Applied Linguistics
Faculty of Humanities
University of Johannesburg, South Africa
ORCID: https://orcid.org/0000- 0002-1201-4341

Dr Peter Babajide Oloba is a lecturer and campus coordinator in the Department of Languages, Cultural Studies, and Applied Linguistics at the University of Johannesburg, South Africa. He holds a PhD, an MEd (with distinction), and a BEd Honours in Educational Leadership and Management from the University of Johannesburg, as well as a Bachelor of Education in Guidance and Counselling and a Nigerian Certificate in Education. His research focuses on educational leadership, systemic change, digital pedagogy, curriculum development, teacher well-being, and social justice in education. Dr Oloba has published extensively in DHET-accredited journals, presented at local and international conferences, and actively supervises Honours, Master's, and PhD students, contributing to transformative scholarship within African and global education contexts.

**Adelaide Rethabile Motshabi Pitso-Mbili**
Department of Tourism, Hospitality and Sport Management
Faculty of Management and Public Administration Sciences
Walter Sisulu University, South Africa
ORCID: https://orcid.org/0000-0003-3878-8641

Dr Adelaide Pitso-Mbili is a dedicated and passionate academic specialising in hospitality and tourism management. She holds a PhD, Master's and a BTech in Management Sciences specialising in Hospitality and Tourism Management from

Durban University of Technology. She also holds a National Diploma in Hospitality Management from the University of Johannesburg. Furthermore, she holds a Postgraduate Certificate in Education specialising in Senior Phase and Further Education and Training from the University of South Africa. She recently obtained her Postgraduate Diploma in Higher Education at Rhodes University. Her research interests include emerging technologies in hospitality and tourism education, technology-enhanced learning, sustainability in hospitality education, and innovative approaches to teaching practical subjects. She is interested in integrating digital tools and connectivist approaches to support teaching and learning in consumer studies, hospitality, and tourism within higher education contexts.

**Memory Ranga**
Department of Business and Application Development,
Faculty of Engineering, Built Environment and Information Technology,
Walter Sisulu University, East London, South Africa
ORCID: https://orcid.org/0000-0002-8970-4789

Dr Memory Ranga is a senior lecturer in the Department of Business and Application Development at Walter Sisulu University, South Africa. Ranga holds a PhD in Information and Technology, specialising in IT governance. Her research focuses on IT governance and ICT for development. She is a member of the ICT for Sustainable Development Research Niche Area at Walter Sisulu University. Ranga is a certified information systems auditor, a certified change management practitioner, and TOGAF (The Open Group Architecture Framework) certified. She serves as the internal editor for books. Leveraging her proven expertise and hands-on experience in executing security governance, conducting risk assessments, ensuring compliance, performing security audits, assessing the effectiveness of controls, managing ISMS, ensuring data privacy, responding to incidents, and overseeing cyber operations, Ranga successfully provides consultation for large enterprise security projects within various government departments in South Africa. In her efforts to raise awareness of cybersecurity, AI, and ICT governance, Ranga has dedicated time to sharing her experience and knowledge with the government sector and academic institutions.

**Petronella Saal**
Equitable Economies and Education (EEE)
Human Sciences Research Council (HSRC), South Africa
ORCID: https://orcid.org/0000-0003-3338-5392

Dr Saal is a Research Specialist in the Equitable Economies and Education unit at the Human Sciences Research Council, South Africa. She completed a doctorate in Computer-Integrated Education with honorary colours at the University of Pretoria (2020). She also holds an honours degree in Education Management, Law, and Policy from the University of Pretoria (2015). Saal specialises in Computer-Integrated Education in primary and secondary schools. Her research interests are comparative education, cross-sectional studies, educational technology, mathematics and science education, and artificial intelligence (AI). She has led and contributed to projects on AI in education, digital game-based learning, and AI for governance and accountability.

**Winter Sinkala**
Department of Mathematical Sciences and Computing
Faculty of Natural Sciences
Walter Sisulu University, Mthatha Campus, South Africa
ORCID: https://orcid.org/0000-0002-2263-7286

Prof. Sinkala is an Associate Professor of Applied Mathematics in the Department of Mathematical Sciences and Computing at Walter Sisulu University, South Africa. He holds a PhD in Mathematics from the University of KwaZulu-Natal and specialises in Lie Symmetry Analysis of Differential Equations. His research portfolio is structured around several key methodological themes within symmetry analysis, including innovative approaches to finding Lie point and non-classical symmetries, group classification of partial differential equations, construction of conservation laws, symmetry reductions via classical and double reduction methods leading to invariant solutions, integration of systems of first-order ordinary differential equations, and the construction of equivalence transformations to relate equivalent differential equations. He also has strong scholarly interests in Mathematics Education and computational fixed-point approaches. Professor Sinkala is actively involved in postgraduate supervision and academic leadership and has contributed to institutional quality assurance and research development initiatives.

**Nompilo Tshuma**
Centre for Higher and Adult Education
Department of Curriculum Studies
Faculty of Education, Stellenbosch University, South Africa
ORCID: https://orcid.org/0000-0002-7842-5426

Dr Tshuma is a senior lecturer and researcher interested in critical and nuanced explorations of educational technology integration in higher education. Her PhD in Information Systems, obtained at Rhodes University, South Africa, set her on the critical path, allowing her to unpack the socio-political nature of technological systems within higher education and the resistance responses that these engendered.

## RESEARCH JUSTIFICATION

Postgraduate supervision in higher education is undergoing a period of significant transformation, driven by the rapid integration of generative artificial intelligence into research and academic writing practices. While AI tools are increasingly embedded in the methods by which postgraduate students search for literature, draft manuscripts, analyse data, and seek feedback, supervisory models have not evolved at the same pace. Existing scholarship tends to focus broadly on AI in teaching and learning, often overlooking the distinctive pedagogical, ethical, and epistemological dynamics inherent in postgraduate supervision. As highlighted in the book, there is a limited yet growing body of research examining the quality of supervision in AI-rich environments; however, this evidence remains fragmented and under-theorised. This gap presents a pressing institutional challenge: universities are expected to maintain rigorous research standards while navigating technological disruption without a sufficiently developed framework for AI-informed supervision. The volume thus represents a necessary scholarly intervention that consolidates emerging insights and reframes supervision as a critical site of inquiry in the age of artificial intelligence.

A second and equally significant justification lies in the limited attention given to supervisors' preparedness, digital competence, and institutional support systems in relation to AI integration. Much of the current discourse concentrates on student use, academic integrity concerns, or technological innovation, while insufficiently interrogating the supervisory role as both mentor and gatekeeper of research quality. The volume identifies that supervisors often operate within environments where policy clarity, structured training, and practical guidance on AI use are inadequate. This is particularly pronounced in resource-constrained contexts, where digital transformation is uneven and access to AI literacy development is restricted. Without deliberate capacity building, these disparities are likely to widen inequities across institutions and disciplines. The volume addresses this lacuna by foregrounding supervisors' agency, professional development needs, and ethical responsibilities, thereby repositioning them as central actors in shaping responsible AI engagement within postgraduate research.

Beyond structural readiness, the book is substantiated by the pedagogical and developmental tensions introduced by AI-mediated scholarship. Generative tools can enhance efficiency and linguistic fluency; however, they also pose risks to the cultivation of deep analytical thinking, methodological rigour, and scholarly voice. Evidence presented in the volume indicates concerns regarding superficially polished outputs masking conceptual weaknesses, reduced confidence in independent academic writing, and emerging patterns of over-reliance that may undermine intellectual growth. These dynamics directly challenge the foundational purpose of postgraduate supervision, which is not merely to produce a thesis but to develop autonomous researchers capable of critical, original knowledge production. The book thus transcends technological optimism or alarmism by interrogating how supervision can remain relational, dialogical, and human-centred in an AI-enhanced academic landscape. This nuanced positioning underscores the need for a framework that safeguards academic integrity while embracing innovation.

Ultimately, this book serves as a timely, integrative contribution that synthesises contemporary scholarship and translates it into actionable insights for supervisors, institutions, and policymakers. By addressing ethics, governance, fairness, AI literacy, policy development, and integrity within postgraduate contexts, the volume offers a coherent roadmap for reimagining supervision in a technologically evolving academy. It bridges theoretical discourse and practical application, fostering dialogue between researchers and practitioners while situating AI within broader concerns of equity, quality assurance, and scholarly formation. In doing so, the book not only fills a significant gap in the literature but also provides a strategic foundation for safeguarding the intellectual and ethical standards of postgraduate research amid accelerating digital transformation.

# PREFACE

This edited volume emerged from a pressing and practical question confronting universities worldwide: what constitutes responsible, high-quality postgraduate supervision in an era where generative and analytic AI tools are embedded throughout the research lifecycle? Within a remarkably short period, artificial intelligence has transitioned from peripheral experimentation to an integral aspect of everyday academic practice, influencing how postgraduate candidates conduct literature searches, structure arguments, analyse data, draft chapters, and refine language. However, supervision has never been defined by speed or technical efficiency. At its core, it is an intellectual apprenticeship grounded in mentorship, ethical stewardship, scholarly dialogue, and the development of independent researchers. This book was conceived from the recognition that technological acceleration demands not reactive anxiety, but rather careful, principled reflection on how supervision can remain pedagogically coherent and academically rigorous in the presence of AI.

The call for contributions invited scholars to interrogate the intersection of AI and postgraduate supervision, with particular attention to critical thinking, creativity, intellectual curiosity, and academic integrity. These concerns are not incidental; they represent the enduring purposes of postgraduate research. While AI can enhance productivity and provide valuable cognitive support, it can also foster superficial synthesis, uncritical reliance on machine-generated outputs, and the subtle erosion of scholarly voice if left unchecked. The chapters collected in this volume therefore share a common commitment: to clarify what must remain human-centered in supervision and to explore how AI can be harnessed in ways that strengthen, rather than dilute, intellectual development. Contributors engage with both practical supervisory moments—proposal formulation, literature review construction, methodological decision-making, writing feedback, examination preparation—and broader institutional questions of governance, policy, equity, and academic integrity. Importantly, the volume acknowledges that the implications of AI are not uniform; they vary across disciplines, institutional capacities, and global contexts, including settings in the Global South where digital transformation presents both opportunities and constraints.

Ultimately, this book is presented as a set of interconnected conversations rather than a single prescriptive model. Across the chapters, a recurring insight emerges: AI has intensified the need for explicit supervisory design. Expectations regarding drafting, citation practices, methodological reasoning, transparency of tool use, and intellectual ownership can no longer be implicit. Supervisors and candidates must negotiate shared understandings about acceptable assistance, responsible authorship, and the cultivation of independent thought. Institutions, in turn, must move beyond punitive responses towards educative, capacity-building frameworks that promote ethical AI literacy. This volume is intended for supervisors, postgraduate candidates, academic developers, and institutional leaders seeking thoughtful engagement with these evolving challenges. Its central message is clear: artificial intelligence should neither be feared nor uncritically embraced but situated within the enduring commitments of postgraduate education—rigour, integrity, creativity, equity, and the development of independent scholars.

# FOREWORD

The incorporation of generative artificial intelligence into higher education marks one of the most significant intellectual transformations of our era. While much public and academic discourse has centred on teaching, evaluation, and academic honesty, relatively little attention has been given to postgraduate supervision - the domain where scholars are nurtured, research identities are formed, and disciplinary knowledge is advanced. This collection addresses this void with precision and urgency. It recognises that supervision is not merely a technical procedure concerned with outcomes, but a relational, ethical, and intellectual practice that moulds the future of scholarship itself.

Artificial intelligence now shapes how postgraduate students explore literature, construct arguments, analyse data, and refine their writing. These advances offer undeniable efficiency, yet they also raise profound pedagogical and epistemological issues: How can supervisors foster originality, critical depth, and methodological rigour when generative AI tools can produce fluent, seemingly sophisticated text within seconds? How should institutions promote responsible AI engagement without impeding innovation? This book tackles these questions head-on, offering a balanced, principled response that steers clear of both technological alarmism and uncritical enthusiasm.

A distinguishing strength of this collection lies in its emphasis on supervisors as key figures in this transition. It underscores the need for institutional transparency, ethical guidance, and professional development to ensure that supervision remains grounded in scholarly integrity and intellectual rigour. At the same time, it acknowledges broader concerns about equity, digital accessibility, and fairness, particularly across diverse institutional settings. By embedding AI within broader discussions on governance, academic responsibility, and researcher advancement, the book provides not only analysis but also guidance.

This is a timely and essential contribution. It reaffirms that, even in a period of rapid technological evolution, the fundamental goal of postgraduate supervision remains the nurturing of independent, critically engaged scholars capable of generating original and ethically sound knowledge.

**Prof. D.C. Geduld**

Nelson Mandela University, South Africa

## ACKNOWLEDGEMENTS

We express our sincere gratitude to the many colleagues and contributors whose scholarship and collegial engagement made this volume possible. An edited collection is, by its nature, a collaborative endeavour—it depends not only on the quality of individual chapters but also on the shared commitment of authors to editorial processes, timelines, revisions, and the collective shaping of a coherent book. We thank all chapter contributors for their intellectual generosity, professionalism, and willingness to refine their work in response to editorial guidance. We are mindful that the editorial process extended well beyond our initial timelines, and we do not take for granted the patience and steadfast commitment of the contributors throughout this odyssey.

We also acknowledge, with appreciation, colleagues who sustained us with their genuine advisory support throughout this project's duration. We value their confidence in the book's scholarly aims and their commitment to enabling research and publication initiatives that advance postgraduate education.

We thank the Research and Innovation Directorate and the Deputy Vice-Chancellor for Research and Internationalisation at Walter Sisulu University for their support during the project.

Finally, we acknowledge the reviewers and the broader institutional and scholarly communities that continue to sustain postgraduate supervision as a space of rigorous inquiry and intellectual formation. It is our hope that this volume will serve that community by contributing thoughtfully to current debates and practices at the intersection of artificial intelligence and postgraduate education.

# TABLE OF CONTENTS

COPYRIGHT AND LICENSING ........ i
DISCLAIMER ........ i
PEER REVIEW CONFIRMATION ........ i
PUBLISHER'S INFORMATION ........ ii
EDITORS' PROFILE ........ iii
RESEARCH JUSTIFICATION ........ xv
PREFACE ........ xvii
FOREWORD ........ xix
ACKNOWLEDGEMENTS ........ xx
Role of Generative Artificial Intelligence in Transforming Supervision Dynamics in Postgraduate Education: A Systematic Literature Review ........ - 1 -
Generative AI as a 'Precipitant' of Challenges in Doctoral Supervision: A Dialogue Among Supervisors ........ - 17 -
Redefining Postgraduate Supervision in the Age of AI: Balancing Technology and Human Mentorship ........ - 38 -
Artificial Intelligence and the PhD: Navigating Doctoralness in the Digital Age ... - 55 -
Exploring Supervisors' Readiness to Integrate AI Tools in Postgraduate Supervision in Higher Education ........ - 74 -
Supervision Skills of Supervisors in AI-enhanced Environments: Perspectives on Postgraduate Supervision ........ - 93 -
Challenges and Risks of Utilising AI Tools for Postgraduate Supervision in Higher Education: Perspectives from South Africa ........ - 111 -
Potential AI-based Use of Fuzzy Cognitive Mapping in Postgraduate Supervision in Higher Education ........ - 130 -
Fostering Intellectual Autonomy in AI-Mediated Postgraduate Supervision: A Gamified Approach ........ - 146 -

Sustainable Mentorship Practices: Designing Frameworks That Blend AI Efficiencies with Human Intuition ........................................................................................ - 166 -

Transforming Research Supervision: Ethical and Literacy Imperatives in the Era of Generative AI ........................................................................................ - 191 -

Assessing Research Integrity in the Age of AI: A Longitudinal Analysis Using an AI Misuse Impact Index ........................................................................................ - 218 -

Integrating Artificial Intelligence in Postgraduate Supervision: Emergent Opportunities, Challenges, and Strategic Responses for Institutions .......................................... - 244 -

Effective AI Integration in Postgraduate Supervision Practices: Policy Implications for South Africa ........................................................................................ - 263 -

CHAPTER ONE

# Role of Generative Artificial Intelligence in Transforming Supervision Dynamics in Postgraduate Education: A Systematic Literature Review


**Godwin P. Dzvapatsva[1]**
**Edmore Chinhamo[2]**
**Courage Matobobo[3]**
**Memory M. Ranga[4]**
**Prince D. N. Ncube[5]**

AFFILIATIONS

[1]Department of Computing, University of Suffolk, United Kingdom.
[2]Executive Business Centre, Greenwich Business School, University of Greenwich, London, United Kingdom.
[3&4]Business and Applications Development, Walter Sisulu University, Buffalo City, South Africa.
[5]Networking and IT Support, Walter Sisulu University, Buffalo City, South Africa.

CORRESPONDENCE

Email: g.dzvapatsva@uos.ac.uk



REFERENCE

Dzvapatsva, G. P., Chinhamo, E., Matobobo, C., Ranga, M. M., & Ncube, P. D. N. (2026). Role of generative artificial intelligence in transforming supervision dynamics in postgraduate education: A systematic literature review. In BI. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 1–16). ERRCD Forum. https://doi.org/10.38140/obp4-2026-01





**Abstract:** Postgraduate supervision plays a critical role in shaping research outcomes, student development, and the mentor-mentee relationship. However, traditional supervision practices, often characterised by limited flexibility and heavy reliance on supervisors, can constrain student growth. The emergence of GenAI presents new opportunities for personalised guidance, faster communication, and increased student autonomy. This study explores the role of GenAI in transforming mentor-mentee relationships, identifying potential benefits and implications for postgraduate education. Adopting a qualitative approach, this study conducted a PRISMA-guided systematic review of relevant literature across Scopus, Web of Science, IEEE Xplore, ScienceDirect, Springer, and Google Scholar. The findings indicate that GenAI enhances supervision by improving feedback and critical thinking, promoting student autonomy and motivation, and introducing considerations for ethical and academic integrity. Effective implementation of GenAI in postgraduate education requires a balanced approach that leverages technological advancements while preserving the relational and empathetic aspects of mentor-mentee interactions. Overall, this study underscores the need for further research to investigate the long-term effects of GenAI on academic supervision and to establish best practices for integrating AI tools that enhance, rather than undermine, the mentorship experience. The study relied on secondary data, and future studies should focus on collecting primary data on the role of artificial intelligence in the mentor-mentee relationship.



***Keywords:*** Generative-artificial intelligence, higher education, mentor-mentee, postgraduate supervision, technology-enhanced supervision.


## 1. Introduction

Artificial Intelligence (AI) has become a prominent topic in public discourse, with growing discussions about its transformative potential. AI is increasingly making its presence felt across various sectors of the economy, such as health, finance, retail, manufacturing, and education, among others. In the context of this study, we focus on generative artificial intelligence tools, which are a subset of AI. Generative artificial intelligence (GenAI) is transforming educational practices by reshaping how students access and engage with learning content (Geerling et al., 2023). GenAI provides students with more immediate access to information and learning resources, which may contribute to improved learning outcomes. However, if educational assessments are not designed to account for GenAI tools, there is a risk that students may attain academic qualifications without demonstrating genuine understanding or critical thinking.

At the postgraduate level, academics' interest in discussions surrounding GenAI is rapidly increasing due to its ability to provide personalised learning experiences for students and track learner output. Moreover, GenAI has enabled students to conduct extensive research amidst concerns over compromised academic integrity. Students' approaches to writing formative and summative assignments have undergone significant changes due to GenAI-powered text creation tools, necessitating that academic supervisors understand how to effectively integrate these tools into the management of traditional assignment responsibilities (Duvignau, 2024). GenAI is increasingly transforming postgraduate supervision and reshaping the mentor–mentee relationship. Literature highlights improvements in communication, personalised guidance, dialogue quality (Lewis & Clutterbuck, 2019), shortened supervision timelines (George, 2023; Vos & Armstrong, 2019), efficient task administration, supervisor diligence, increased autonomy, and enhanced ethical monitoring. These developments underscore the evolving role of AI in contemporary supervision practices.

Postgraduate students undertake research studies under the supervision of one mentor or a research team. In traditional supervision models, limited tools foster over-reliance on supervisor guidance, restricting independent problem-solving. These arrangements are frequently marked by long turnaround times for feedback, poor planning, and occasionally strained mentor-mentee interactions (Dai et al., 2023; Paulsen & Schmidt-Crawford, 2017). Although this is the case, some authors, such as Bouzar et al. (2025), argue that traditional supervisory models are invaluable for their engagement and contextual relevance as a role model in physical space for the mentee. Academic sources (Lewis & Clutterbuck, 2019; Bearman, Boud, & Konradsen, 2025) report that AI has the potential to support postgraduate supervision and that GenAI-generated supervisor feedback can be easily accessible and understandable. However, empirical studies proving that supervisors have used it are scarce (Thong et al., 2025). Despite this scarcity, there is overwhelming evidence of students using AI for their assignments, with most universities worldwide significantly affected (Duvignau, 2024). Consequently, managing the quality of research output becomes problematic (Bjelobaba et al., 2024; Chauhan & Currie, 2024). While AI offers students greater autonomy in defining research directions and preparing initial drafts, it can also undermine critical thinking and originality (Aymen & Zakarya, 2024). This challenge places a burden on supervisors to support mentees in almost equal measure to the traditional supervision model.

### 1.1 Problem statement

In an ideal postgraduate supervision environment, supervision is characterised by regular, meaningful interaction between supervisors and students, timely feedback, mutual trust, and sustained academic mentorship that supports both scholarly development and personal growth. However, this ideal has become increasingly difficult to maintain in many higher education institutions. Growing postgraduate enrolments have significantly increased supervisors' workloads, placing additional administrative and academic demands on them that are manageable with smaller cohorts. At the same time, the shortage of qualified and available mentors has resulted in extended supervision periods, further straining the supervisory process and challenging the effectiveness of traditional supervision models (Chapman et al., 2021; Gallacher, 1997; Malik & Malik, 2015). The increasing demand to revamp traditional supervision models calls for more dynamic and flexible approaches that prioritise student autonomy, strengthen mentor–mentee relationships, and provide continuous feedback (Kimani, 2014). Despite extensive literature on conventional postgraduate supervision, there is a lack of research examining how GenAI can enhance the postgraduate supervisory experience. Innovative ways to improve the effectiveness of mentoring and supervision are becoming increasingly necessary as postgraduate education grows more challenging and complex. By providing tools and systems that can revolutionise the way mentoring and supervision are conducted, GenAI holds promise for addressing these issues. However, there are also significant concerns about how GenAI will affect the mentor-mentee relationship, supervisory dynamics, and ethical issues when incorporated into postgraduate education (Köbis & Mehner, 2021).

Based on the highlighted problem, the study is guided by the following research questions:

i. How can GenAI be used to transform mentor-mentee relationships and supervision dynamics in postgraduate education?
ii. What institutional, cultural, and ethical factors influence the effective and responsible use of GenAI in postgraduate supervision?

## 2. Methodology

This study employed a qualitative systematic literature review (SLR) to critically investigate the evolving role of GenAI in shaping mentor-mentee relationships and supervision dynamics within postgraduate education. The selection of an SLR was informed by its capacity to facilitate a comprehensive, transparent, and reproducible synthesis of peer-reviewed scholarly evidence, particularly relevant in a rapidly advancing and multidisciplinary field such as GenAI in education. The review adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines to ensure methodological rigour and compliance with best practices in evidence synthesis. This methodological framework guided the structured identification, screening, and inclusion of relevant literature, ensuring analytical coherence in mapping the conceptual and thematic contours of GenAI's impact on postgraduate supervision. By situating the review within a rigorous and transparent framework, the study seeks to provide substantive insights into the ways in which GenAI technologies reshape traditional supervisory models and pedagogical relationships. Data were collected from six (6) databases: Scopus, Web of Science, IEEE Xplore, ScienceDirect, Springer, and Google Scholar. These databases were selected for this scoping review due to their extensive coverage of high-quality, peer-reviewed literature in the fields of education, technology, and artificial intelligence. The term 'artificial intelligence' was utilised as part of our search string instead of 'Generative AI' to avoid constraining the retrieved articles. Table 1 presents a list of the search strings employed in this study.

***Table 1:*** *Search Strings used for Document Identification*

| **Database** | **Search String** |
|---|---|
| Scopus | ( "artificial intelligence" OR "machine learning" OR "deep learning" OR "AI-driven" OR "intelligent tutoring system" OR "chatbot" OR "automated feedback" OR "predictive analytics" ) AND ( "mentorship" OR "mentor-mentee relationship" OR "academic supervision" OR "postgraduate supervision" OR "graduate advising" OR "doctoral supervision" OR "PhD mentorship" ) AND ( "higher education" OR "postgraduate education" OR "graduate studies" OR "doctoral education" OR "university learning environments" OR "supervision dynamics" ) |
| Web of Science | ( "artificial intelligence" OR "machine learning" OR "deep learning" OR "AI-driven" OR "intelligent tutoring system" OR "chatbot" OR "automated feedback" OR "predictive analytics" ) AND ( "mentorship" OR "mentor-mentee relationship" OR "academic supervision" OR "postgraduate supervision" OR "graduate advising" OR "doctoral supervision" OR "PhD mentorship" ) AND ( "higher education" OR "postgraduate education" OR "graduate studies" OR "doctoral education" OR "university learning environments" OR "supervision dynamics" ) |
| IEEE Xplore | postgraduate supervision and education |
| ScienceDirect | ("artificial intelligence") AND ("supervision mentorship" OR "postgraduate mentorship" OR "academic supervision" OR "postgraduate supervision") AND ("higher education" OR "graduate studies" OR "doctoral education" OR "university learning") |
| Springer | ("artificial intelligence") AND ("mentor-mentee relationship" OR "academic supervision" OR "postgraduate supervision" OR "graduate advising" OR "doctoral supervision" OR "PhD mentorship") AND ("higher education" OR "postgraduate education" OR "graduate studies" OR "doctoral education" OR "university learning environments" OR "supervision dynamics") |
| Google Scholar | ("artificial intelligence") AND ("postgraduate supervision" OR "graduate advising") AND ("higher education" OR "postgraduate education" OR "graduate studies" OR "university learning environments" OR "supervision dynamics") |

Search strings were generated for each database. The initial database search retrieved 493 records from the six databases, as illustrated in Figure 1. Among the 493 records identified, there was an overlap, as the same records were retrieved from different databases. Overall, 28 duplicate records were excluded, leaving 465 records to be screened. We used the article titles and abstracts to further screen the records based on the eligibility criteria outlined below. Four hundred and fifty (450) records were excluded as they met the exclusion criteria, leaving us with 15 records. We then successfully retrieved all 15 records and reviewed the full texts of each to determine whether they met the inclusion criteria. A further two records were excluded, resulting in a total of 13 articles to be included in the review. To increase the number of studies included in the analysis, additional searches were conducted using the Google search engine, from which three further studies were identified and added. In total, sixteen studies were included in the final analysis, as they met the study's inclusion criteria.

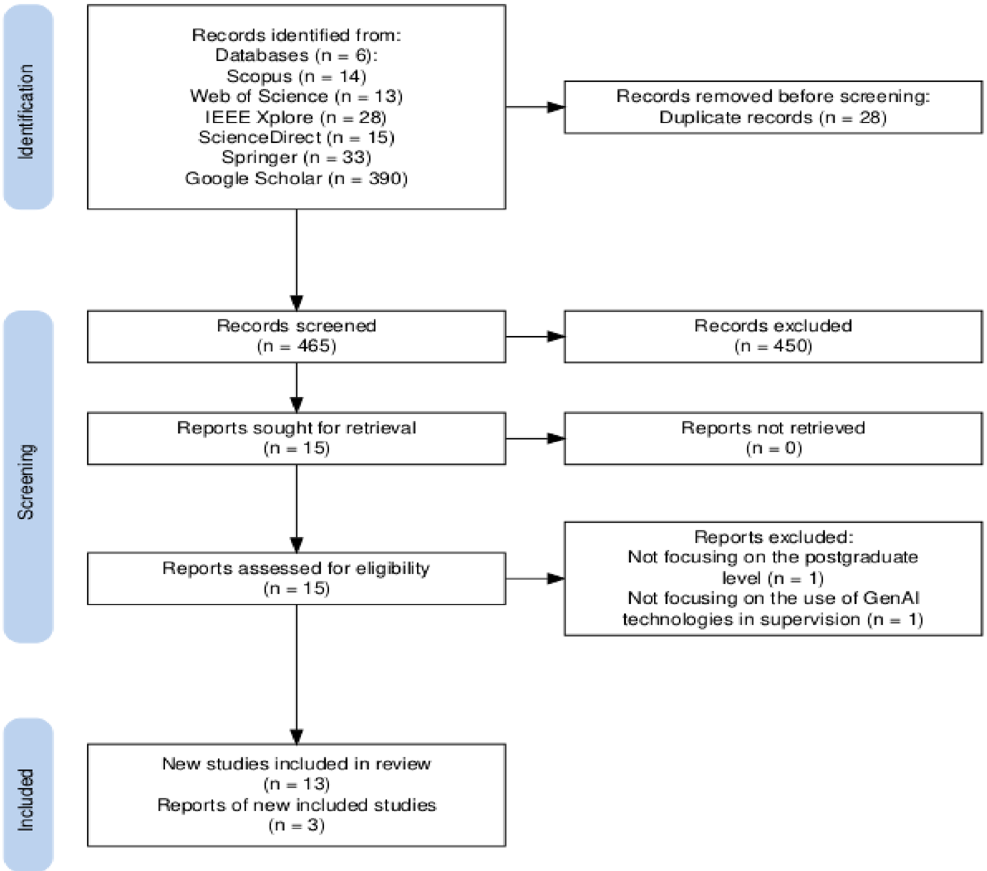



***Figure 1:*** *PRISMA flow diagram of the study's screening process*

NB: The search string used to identify documents from IEEE Xplore was broadened to retrieve more documents, as the initial search string retrieved only one paper from the databases.

Eligibility for inclusion in this study was determined based on a clear set of scope-related considerations. Only full-text, peer-reviewed studies written in English were considered. The review specifically focused on studies examining the use of generative AI technologies within the context of postgraduate supervision. To ensure relevance, included studies had to concentrate on the postgraduate level and explicitly address aspects of the mentor-mentee relationship and/or the supervision process. Studies that did not meet these criteria—such as those unrelated to generative AI, not focused on postgraduate education, not addressing supervision or mentoring, or not published in English—were excluded from the review. The documents selected for analysis in this study are listed in Tables 2a and 2b.

***Table 2a:** Documents analysed*

| No. | Authors | Year | Title | Study Design | Types of AI |
|---|---|---|---|---|---|
| 1 | Wright, A | 2024 | Postgraduate Supervision in a ChatGPT World: What's Next? | Qualitative literature review | GenAI |
| 2 | Jensen, L.X., Bearman, M., Boud, D. and Konradsen, F | 2025 | Feedback encounters in doctoral supervision: the role of generative AI chatbots | Innovative research | GenAI chatbot |
| 3 | Chang, C.N., Hui, J., Justus-Smith, C. and Wang, T.W | 2024 | Navigating STEM careers with AI mentors: a new IDP journey | Mixed methods | GenAI |
| 4 | Thong, C.L., Atallah, Z., Islam, S., Lim, W. and Cherukuri, A. K | 2025 | AI-Powered Tools for Doctoral Supervision in Higher Education: A Systematic Review | Systematic Review | GenAI, LLM, NLP, ITS, AI-powered chatbots, AI for assessments, Predictive modelling, and learning analytics |
| 5 | Dai, Y., Lai, S., Lim, C.P. and Liu, A | 2023 | ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students | Qualitative | GenAI |
| 6 | Bouzar, A., El Idrissi, K., Ghourdou, T. and Ali, N | 2025 | Supervisory Feedback vs. AI: A Comparative Study on Postgraduate Student Satisfaction | Cross-sectional | GenAI |
| 7 | Harding, D. and Boyd, P | 2024 | Generative AI and PHD supervision: a covert third wheel or a seat at the table? | Qualitative exploration | GenAI |
| 8 | Mhlanga, D. and Ndhlovu, E | 2024 | Digital transformation in higher education and postgraduate research supervision in Africa: A critique of 4IR-based interventions in open distance education | Critical document analysis | 4IR |
| 9 | Ndjuluwa, L., Adebisi, J.A. and Abdulsalam, K.A | 2024 | A Review on Digital Tools for Engineering Postgraduate Education in post-Covid Era | Systematic literature review | Digital tools |
| 10 | Iwashokun, O. and Ade-Ibijola, A | 2022 | Structural Vetting of Research Proposals: Problematisation and Solving with Artificial Intelligence | Mixed methods | Grammarly, ProWritingAid, Paperrater, Typely, Wordy, Chatbots, ITS, VR, serious games, genAI in education |

***Table 2b:** Documents Analysed*

| No. | Authors | Year | Title | Study Design | Types of AI |
|---|---|---|---|---|---|
| 11 | Serek, A. and Zhaparov, M | 2024 | Optimising preference satisfaction with a genetic algorithm in matching students to supervisors | Experimental | Genetic algorithm |

| | | | | | |
|---|---|---|---|---|---|
| 12 | Juma, M. N | 2024 | Navigating the ChatGPT Theological Terrain: Considerations for Graduate Theology Students | Qualitative literature review | GenAI |
| 13 | Sim, K.N., Northcote, M. and Lim, C. P | 2023 | Technology-enabled undergraduate and postgraduate research supervision | Conceptual thematic review | ChatGPT and language models |
| 14 | Mbodia, M | 2025 | Postgraduate Supervision in the Age of ChatGPT: Redefining the Role of Supervisors | Literature review | GenAI |
| 15 | Boyd, P. and Harding, D | 2025 | Generative AI: reconfiguring supervision and doctoral research | Mixed methods | GenAI |
| 16 | Iatrellis, O., Bania, A., Samaras, N., Kosmopoulou, I. and Panagiotakopoulos, T | 2025 | ChatGPT in doctoral supervision: Proposing a tripartite mentoring model for AI-assisted academic guidance | Structured evaluation | GenAI |

Thematic analysis was employed in this study. Atlas.ti was used for analysis purposes to create codes and themes, following the six steps outlined by Clarke and Braun (2026). The steps included: familiarisation, coding, generating themes, reviewing themes, defining and naming themes, and writing up.

## 3. Presentation of Results

The study followed Braun and Clarke's (2006) six steps for analysing qualitative data thematically. The results are presented as themes, as shown in Table 3. Activity Theory was employed as the lens to guide the investigation, as it provides a comprehensive framework for understanding human activity systems, including the interactions between individuals, tools, rules, and the community. This approach allows for a deeper analysis of how these elements influence and shape the focus of the study.

***Table 3:*** *Themes and codes*

| **Theme** | **Codes** |
|---|---|
| GenAI as a complementary tool in supervision | GenAI as a complementary tool<br>Complementary role of GenAI<br>GenAI as a supervisor<br>Hybrid supervision model<br>GenAI for supervision efficiency |
| Ethical and academic integrity concerns | Ethical ambiguity<br>Ethical concerns<br>GenAI and academic integrity<br>Privacy issues<br>Misinformation risk |
| Shifting roles of supervisors and mentees | Evolving role of the supervisor<br>Changing supervisor roles<br>Mentor role redefinition<br>Student agency |

| | |
|---|---|
| GenAI-enhanced feedback and critical thinking | GenAI clarity vs supervisor clarity<br>Formative value of feedback<br>Critical thinking aid<br>Engagement in human supervision |
| Institutional and cultural readiness | Need for institutional guidance<br>Institutional support<br>Cultural and contextual influences<br>Resource constraints<br>GenAI literacy needs |
| Student autonomy and motivation | Autonomy boost<br>Dependence vs critical thinking<br>Overreliance concerns<br>Agency and control<br>Student-driven ICT use |
| Technological and pedagogical integration | Technology-enhanced supervision<br>Pedagogy of supervision<br>Education 4.0 and digital transformation<br>GenAI in postgraduate supervision<br>Predictive algorithms |

### 3.1 GenAI as a complementary tool in supervision

This theme explores the supportive role of artificial intelligence in postgraduate supervision, emphasising its function as a complementary tool rather than a replacement. GenAI enhances supervision by increasing efficiency and addressing shortcomings in traditional models without displacing the human supervisor (Jensen, Bearman, Boud & Konradsen, 2025; Boyd & Harding, 2025; Iatrellis et al., 2025). It assists by automating repetitive tasks such as structural feedback, grammar checks, and initial content vetting, thereby reducing supervisors' workload (Iwashokun & Ade-Ibijola, 2022; Serek & Zhaparov, 2024). GenAI also aids in brainstorming, literature reviews, drafting, clarifying concepts, and synthesising data (Mbodia, 2025; Boyd & Harding, 2025). This enables supervisors to dedicate more time to intellectual guidance and emotional support, fostering a more balanced and human-centred supervisory relationship (Dai et al., 2023). AI-driven platforms like ChatGPT thus contribute to a hybrid supervision model, where GenAI handles routine academic support while human supervisors retain critical mentoring. Additionally, GenAI can help bridge gaps in supervision quality, especially where students may not receive consistent or timely guidance, offering accessible mentorship and academic reinforcement (Harding & Boyd, 2024). Supervisors have the responsibility of guiding students to use AI critically and selectively (Mbodia, 2025).

### 3.2 Ethical and academic integrity concerns

This theme highlights the complex ethical dilemmas surrounding the integration of GenAI into postgraduate supervision, particularly regarding plagiarism, authorship, false citation, data security, risks of over-reliance, authenticity of voice, and unacknowledged use, all of which influence academic integrity (Mbodia, 2025; Boyd & Harding, 2025; Iatrellis et al., 2025). Wright (2024) notes ongoing uncertainty about how AI-generated content should be governed within educational settings. Dai et al. (2023) caution against over-reliance on GenAI tools, citing risks such as algorithmic bias and diminished opportunities for authentic learning. Overreliance on AI can reduce intrinsic motivation and independence in problem-solving (Iatrellis et al., 2025). Juma (2024) and Harding and Boyd (2024) warn that GenAI may compromise the integrity of academic work and the pedagogical value of the supervisory process. Additionally, Chang et al. (2024) raise critical concerns about data privacy, information security, and the potential for GenAI to produce inaccurate or misleading outputs, which could misguide both mentors and mentees. These ethical challenges underscore the urgent need for clear institutional policies and guidelines to ensure the responsible and constructive use of GenAI in supervisory contexts (Mbodia, 2025).

### 3.3 Shifting roles of supervisors and mentees

The roles of supervisors and mentees are undergoing a significant transformation due to the integration of GenAI in postgraduate education. The study by Wright (2024), Mbodia (2025), and Boyd and Harding (2025) shows that supervisors' roles are shifting from traditional directive approaches to facilitative, co-learning methods that emphasise GenAI literacy while fostering student autonomy. As technology is integrated into academic practice, there is a change in supervisory roles to support not just academic development, but also relational and reflective mentoring (Dai et al., 2023). This shift reflects a broader trend towards cultivating autonomy and critical thinking, with supervisors acting as collaborators in navigating digital technologies (Chang et al., 2024; Sim et al., 2023). Similarly, mentees are transitioning from passive recipients of knowledge to self-directed, autonomous researchers who utilise GenAI tools, such as ChatGPT, to guide their academic journeys (Dai et al., 2023). Students take on the role of critical evaluators and synthesisers of AI outputs, which is a complementary and defined task (Iatrellis et al., 2025). However, if GenAI is adopted uncritically, it may function as a surrogate mentor, subtly reshaping supervision and, in some cases, marginalising the role of the human supervisor (Boyd & Harding, 2025). This evolving dynamic calls for redefined expectations, responsibilities, and skill sets in the supervision relationship.

### 3.4 GenAI-enhanced feedback and critical thinking

This study explores the evolving role of Generative Artificial Intelligence (GenAI) in providing feedback and fostering critical thinking within the context of postgraduate supervision. GenAI tools, such as ChatGPT, Perplexity, Scite, Elicit, among others, are often perceived as capable of providing clear and immediate feedback, with some mentees finding such feedback to be more comprehensible than that provided by human supervisors (Bouzar et al., 2025; Mbodia, 2025). Furthermore, GenAI assists students in rehearsing responses, clarifying feedback, and building confidence prior to engaging with their supervisors (Boyd & Harding, 2025). Nevertheless, given that GenAI operates as a "devil's advocate" to stimulate critical thinking (Dai et al., 2023), it lacks the pedagogical depth and contextual understanding that human supervisors can offer (Jensen et al., 2025). GenAI feedback is typically task-oriented and focuses on superficial enhancements, whereas supervisors play a pivotal role in aiding students' development of their intellectual scholarly identity and higher-order thinking skills (Jensen et al., 2025; Mbodia, 2025). The findings indicate that while GenAI can enhance the feedback process, it cannot supplant the development, rational guidance, and individualised support that human supervisors provide (Bouzar et al., 2025).

### 3.5 Institutional and cultural readiness

The theme of institutional and cultural readiness emphasises the crucial role of institutional support and cultural context in effectively integrating GenAI into postgraduate supervision. The studies by Bouzar et al. (2025) and Wright (2024) highlighted the need for structured frameworks, clear policies, and training programmes that can guide mentees and supervisors in using GenAI tools ethically and effectively. Mbodia (2025), Iatrellis et al. (2025), and Boyd and Harding (2025) noted a lack of clear guidelines for the use of AI in higher education institutions. According to Sim et al. (2023), institutional context and cultural background significantly impact how ICT and GenAI tools are perceived and used, necessitating local approaches. However, resource limitations, such as limited access to reliable internet and computers, remain significant barriers, especially in disadvantaged institutions (Mhlanga & Ndhlovu, 2024). These disparities in resources create inequalities among university students (Mbodia, 2025). Institution-wide GenAI literacy training fosters confidence, ensures responsible use, and promotes fair participation in AI-enabled academic settings (Dai et al., 2023).

### 3.6 Student autonomy and motivation

The theme of student autonomy and motivation explores the dual roles of GenAI tools, such as ChatGPT, in promoting student motivation and autonomy in postgraduate supervision. ChatGPT provides immediate on-demand support, enhancing student confidence, reducing dependency on supervisor availability, and promoting self-directed learning by minimising the need for supervisors for routine queries (Dai et al., 2023; Boyd & Harding, 2025; Iatrellis et al., 2025). Due to the accessibility of some GenAI tools, students are empowered to take greater control of their academic progress, particularly by allowing them to seek feedback on their own terms, thus enhancing their perceived agency (Jensen et al., 2025). Similarly, Sim et al. (2023) and Mbodia (2025) emphasise the autonomous use of ICT tools, such as ChatGPT and learning management systems (LMS), by students to guide their research and learning processes. However, concerns such as overdependence arise with the growing use of GenAI, which may hinder the development of critical thinking and problem-solving skills (Chang et al., 2024; Wright, 2024).

### 3.7 Technological and pedagogical integration

This theme explores the growing convergence of technology and pedagogy in postgraduate research supervision. Integrating GenAI tools into postgraduate supervision has significantly improved supervisory practices by enabling more flexible, responsive, and accessible mentoring (Sim et al., 2023). The use of digital technologies, such as AI, in education is shaping the way supervision is conducted and redefining mentoring approaches to align with the evolving needs of students. AI-driven tools, such as ChatGPT, are being utilised to support supervision in a technology-driven, student-centred environment, in line with Education 4.0 trends (Iwashokun & Ade-Ibijola, 2022; Leokadia et al., 2024). These tools enhance academic engagement, streamline feedback, and facilitate collaboration across digital platforms. Furthermore, the use of machine learning algorithms in supervision enables personalised academic support by predicting mentees' needs and offering tailored guidance (Thong et al., 2025). The adoption of hybrid supervision models, where AI tools are blended with traditional mentorship, is recommended for successful postgraduate supervision (Mbodia, 2025). However, the integration of GenAI in postgraduate supervision remains ad hoc and often hidden (Boyd & Harding, 2025).

## 4. Discussion of Findings

This chapter explores the role of GenAI in transforming mentor-mentee relationships and supervisory dynamics in postgraduate education. The findings of this study indicate a growing institutional shift from initial resistance to the gradual adoption of these technologies within the sampled contexts. However, the extent to which this shift is widespread across the broader higher education sector remains uneven and context-dependent, suggesting the need for further large-scale and multi-institutional studies to establish the generalisability of this trend. The key findings of this of study were: (1) using GenAI as a complementary tool in supervision; (2) the shifting of supervisor and mentee roles; (3) using GenAI for enhanced feedback and critical thinking; (4) student autonomy and motivation; (5) institutional and cultural readiness; and (6) ethical and academic integrity concerns.

Resonating with social learning concepts, knowledge cannot be separated from the environment in which it was produced. The environment encompasses its people, the tools involved, and the systems in place. The flexibility in communication aligns with the principles of Activity Theory, where tools mediate the relationship between subjects and objects. Evidence consistently demonstrates that mediation tools, especially in the early stages of integrating large language models like ChatGPT, significantly enhance student learning experiences by supporting engagement, comprehension, and personalised learning (Dai et al., 2023; George, 2023; Matobobo et al., 2025). Based on the study's results, it is evident that the use of GenAI in education has expanded the number of mediating artefacts available for research, which in turn impacts the effectiveness of mentoring. Researchers have lamented that while GenAI tools are beneficial for supervision, institutions must ensure that learners have a solid foundation in the fundamental principles of the subject and utilise the complementary aspects of the tools rather than replacing supervisors. For mentees to generate effective prompts, they must possess a foundational understanding of the subject matter. Without sufficient knowledge, there is a greater risk of being misled or producing superficial inputs that lack critical depth. A firm grasp of the topic allows mentees to engage meaningfully with AI tools, enabling them to frame precise and thoughtful queries. Research in education and cognitive science consistently shows that domain knowledge enhances critical thinking and the ability to evaluate information effectively. When used correctly and with appropriate subject knowledge, these tools can significantly amplify the quality of work produced. While some argue that tools like GenAI can fill knowledge gaps (Yagyaeva et al., 2024), they are most effective when guided by an informed user who can evaluate and refine the output.

The role of the mentor is increasingly shifting towards creativity rather than merely addressing the basics, a transition that can be facilitated by GenAI tools (Sim et al., 2023). Unlike traditional supervision, GenAI assists postgraduate students in independently gathering and analysing relevant literature in a short time, allowing the mentee to delve deeper into their selected phenomena of interest. This shift prepares individuals for the role of academic advisors (Harding & Boyd, 2024). Of course, like any tool, over-reliance on GenAI can result in diminished human interaction. Other factors, such as guidance on the flow of document structure and in-depth analysis of results based on environmental factors, still require human experience and insight (Wright, 2024). Furthermore, it is essential to note that, regardless of whether students use GenAI, the supervisor remains responsible and accountable for the final research output produced by the student. Nonetheless, the immediacy of feedback offered by GenAI tools reduces anxiety for the mentee in cases where the mentor delays responses.

Given the rapid increase in the use of these tools, it should be the prerogative of institutions to develop policies and practices that demonstrate readiness for mentor-mentee relationships in a human-AI collaborative environment. Due to the absence of clear implementation and policy frameworks within institutions, our experience reveals a degree of mistrust among practitioners regarding the integrity of the supervision processes facilitated by GenAI. The findings suggest that cultural readiness enables mentors to guide mentees ethically in the use of artificial intelligence in research. It is important to emphasise that awareness of ethical considerations will lead to a focus on process rather than solely on outcomes.

Additionally, the deployment and use of GenAI in research raise ethical concerns, including data privacy and algorithmic bias. As demonstrated by Dai et al. (2023), these issues necessitate the establishment of clear guidelines and ethical standards to govern the use of AI applications in educational settings. This ensures that technology usage aligns with the overarching goals of education and maintains the integrity of mentor-mentee relationships.

The literature lacks discussion on sustainability. There is significant concern regarding the amount of energy required to train these models. Considering green AI, which involves using

AI to transform mentor-mentee relationships and supervision dynamics in postgraduate education, is essential as its usage rapidly increases. This gap highlights the urgent need to align AI innovation with environmentally conscious and educationally impactful objectives. Furthermore, the study's findings are inconclusive regarding whether AI is geared towards personal development or personal performance, as its use in supervision continues to grow.

## 5. Conclusions

This study concludes that Generative Artificial Intelligence (GenAI) possesses significant potential to reshape postgraduate supervision, particularly by influencing the mentor-mentee relationship and supervisory practices within higher education. The reviewed literature consistently indicates that GenAI tools can enhance efficiency in supervision by supporting tasks such as feedback generation, idea exploration, academic writing assistance, and administrative support. These capabilities have the potential to alleviate supervisors' workloads and improve students' access to timely academic guidance. However, the findings also affirm that GenAI cannot replace the uniquely human dimensions of postgraduate supervision. Core supervisory functions, such as providing emotional support, fostering trust, offering contextual and disciplinary judgement, and guiding students' scholarly identity development, remain inherently human. The literature emphasises that effective supervision depends on relational, ethical, and developmental elements that GenAI systems are currently unable to replicate.

Furthermore, the study highlights the necessity for a balanced and complementary integration of GenAI into supervisory practices. While GenAI can augment supervision, unethical or excessive reliance on these tools risks undermining academic integrity, student agency, and the quality of mentorship. Ethical challenges, including bias, transparency, accountability, and data privacy, position supervisors as essential gatekeepers responsible for ensuring the responsible and ethical use of GenAI in postgraduate contexts. Hence, the study concludes that GenAI should be understood as an enabling technology that supports, rather than substitutes for, human mentorship. Its value lies in complementarity rather than replacement, reinforcing the continued centrality of the supervisor in postgraduate education.

## 6. Recommendations and Limitations of the Study

Based on the findings of this study, several recommendations are proposed for higher education institutions, supervisors, and policymakers:

- Higher education institutions should develop clear policies and frameworks that guide the ethical and pedagogically sound utilisation of Generative Artificial Intelligence (GenAI) in postgraduate supervision. These frameworks should define acceptable use, outline responsibilities for both supervisors and students, and address issues related to academic integrity, data protection, and transparency.
- Universities should invest in training programmes for supervisors and postgraduate students to enhance their literacy in GenAI. Such initiatives should focus on developing

a critical awareness of GenAI's capabilities and limitations, thereby enabling users to leverage these tools effectively while maintaining scholarly rigour and ethical standards.

- Supervision models should explicitly foreground the irreplaceable human aspects of mentorship, including emotional support, ethical judgement, and personalised guidance. GenAI should be positioned as a supplementary tool that enhances efficiency, allowing supervisors to dedicate more time to high-level intellectual engagement and relational support.
- Institutions should establish mechanisms for the continuous evaluation of GenAI use in postgraduate supervision. This includes monitoring unintended consequences, addressing emerging ethical concerns, and ensuring that GenAI adoption aligns with sustainability, inclusivity, and student well-being objectives.

From a methodological standpoint, the main limitation of this study lies in its reliance on secondary data. Furthermore, supervision practices vary significantly depending on institutional context and individual supervisory styles, which may affect the generalisability of the findings. Despite these limitations, the study provides a valuable foundation for future research. Subsequent studies should incorporate primary data collection to explore the lived experiences of both mentors and mentees, offering deeper insights into how GenAI can be effectively and responsibly integrated into postgraduate supervision.

## 7. Declarations

**Funding:** This research did not receive any external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Use of Artificial Intelligence:** The current work was supported with the assistance of artificial intelligence technologies (ChatGPT-4 and Grammarly) to assist with refining language for clarity, as confirmed by the authors.

# References

Aymen, D., & Zakarya, B. (2024). *The influence of artificial intelligence on students' critical thinking: A case study of third-year English students at the University Centre of Mila.* Abdelhafid Boussouf University - Mila.

Boyd, P., & Harding, D. (2025). Generative AI: Reconfiguring supervision and doctoral research. *Buildings & Cities, 6*(1), 1–29.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Chang, C. N., Hui, J., Justus-Smith, C., & Wang, T. W. (2024). Navigating STEM careers with AI mentors: A new IDP journey. *Frontiers in Artificial Intelligence, 7*. https://doi.org/10.3389/frai.2024.1461137

Chapman, L., Mysko, C., & Coombridge, H. (2021). Development of teaching, mentoring and supervision skills for basic training registrars: A frustrated apprenticeship? *Internal Medicine Journal, 51*(11), 1847–1853. https://doi.org/10.1111/imj.14935

Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology, 39*(4), 74–88.

Duvignau, R. (2024). ChatGPT has eaten my assignment: A student-centric experiment on supervising writing processes in the AI era. *IEEE Global Engineering Education Conference, EDUCON*. https://doi.org/10.1109/EDUCON60312.2024.10578917

Gallacher, K. (1997). Supervision, mentoring, and coaching. In P. J. Winton, J. A. McCollum, & C. Catlett (Eds.), *Reforming personnel in early intervention* (pp. 191–214). Paul H. Brookes.

Geerling, W., Mateer, G. D., Wooten, J., & Damodaran, N. (2023). ChatGPT has aced the test of understanding in college economics: Now what? *American Economist, 68*(2), 233–245. https://doi.org/10.1177/05694345231169654

George, A. S. (2023). The potential of generative AI to reform graduate education. *Partners Universal International Research Journal, 2*(4), 36–50. https://doi.org/10.5281/zenodo.10421475

Harding, D., & Boyd, P. (2024). Generative AI and PhD supervision: A covert third wheel or a seat at the table? *ARCOM 2024: 40th Annual Conference – Looking Back to Move Forward*. https://ucem.repository.guildhe.ac.uk/id/eprint/125/1/GenAI_registry.pdf

Iatrellis, O., Bania, A., Samaras, N., Kosmopoulou, I., & Panagiotakopoulos, T. (2025). ChatGPT in doctoral supervision: Proposing a tripartite mentoring model for AI-assisted academic guidance. *International Journal of Doctoral Studies, 20*(9), 1-26.

Iwashokun, O., & Ade-Ibijola, A. (2022). Structural vetting of research proposals: Problematisation and solving with artificial intelligence. *Proceedings of the International Conference on Information Systems and Emerging Technologies*. https://ssrn.com/abstract=4331475

Jensen, L. X., Bearman, M., Boud, D., & Konradsen, F. (2025). Feedback encounters in doctoral supervision: The role of generative AI chatbots. *Assessment and Evaluation in Higher Education*. https://doi.org/10.1080/02602938.2025.2478155

Juma, M. N. (2024). Navigating the ChatGPT theological terrain: Considerations for graduate theology students. *Pan-African Journal of Education and Social Sciences*, *5*(2), 158–166. https://doi.org/10.56893/pajes2024v05i02.11

Kimani, E. N. (2014). Challenges in quality control for postgraduate supervision. *International Journal of Humanities Social Sciences and Education (IJHSSE)*, *1*(9), 63–70. http://www.arcjournals.org

Köbis, L., & Mehner, C. (2021). Ethical questions raised by AI-supported mentoring in higher education. *Frontiers in Artificial Intelligence*, *4*. https://doi.org/10.3389/frai.2021.624050

Leokadia, N., Adebisi, J. A., & Khadeejah, A. (2024). A review on digital tools for engineering postgraduate education in the post-Covid era. *Adeleke University Journal of Engineering and Technology*, *7*(1), 222–235. http://www.aujet.adelekeuniversity.edu.ng

Lewis, L., & Clutterbuck, D. (2019). Co-evolution: Exploring synergies between artificial intelligence (AI) and the supervisor. In *Coaching supervision* (pp. 200–216). Routledge.

Malik, S. A., & Malik, S. A. (2015). Graduate school supervisees' relationships with their academic mentors. *Journal of Applied Research in Higher Education*, *7*(2), 211–228. https://doi.org/10.1108/JARHE-07-2013-0032

Matobobo, C., Ncube, P. D. N., Ngesimani, N. L., Dzvapatsva, G. P., & Chinhamo, E. (2025). Enhancing computational thinking and problem-solving in programming education through generative AI: A scoped review. In *2025 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-9). IEEE.

Mbodila, M. (2025). Postgraduate supervision in the age of ChatGPT: Redefining the role of supervisors. *Foundations and Trends® in Information Systems*, *9*(2), 60-68.

Mhlanga, D., & Ndhlovu, E. (2024). Digital transformation in higher education and postgraduate research supervision in Africa: A critique of 4IR-based interventions in open distance education. *Journal of Southwest Jiaotong University*, *59*(5), 1-13. https://doi.org/10.35741/issn.0258-2724.59.5.11

Sim, K. N., Northcote, M., & Ping Lim, C. (2023). Technology-enabled undergraduate and postgraduate research supervision. *Australasian Journal of Educational Technology, 39*(4), 1–5.

Paulsen, T. H., & Schmidt-Crawford, D. A. (2017). Enhancing student teacher supervision through hybridization: Adding e-supervision to the mix. *Journal of Agricultural Education, 58*(2), 166–179. https://doi.org/10.5032/jae.2017.02166

Serek, A., & Zhaparov, M. (2024). Optimizing preference satisfaction with genetic algorithm in matching students to supervisors. *Applied Mathematics and Information Sciences, 18*(1), 133–138. https://doi.org/10.18576/amis/180114

Bouzar, A., El Idrissi, K., Ghourdou, T., & Ali, N. (2025). Supervisory feedback vs. AI: A comparative study on postgraduate student satisfaction. *ISRG Journal of Arts, Humanities & Social Sciences, 3*(1), 465–474. https://doi.org/10.5281/zenodo.14862530

Thong, C. L., Atallah, Z., Islam, S., Lim, W., & Cherukuri, A. K. (2025). AI-powered tools for doctoral supervision in higher education: A systematic review. *Journal of Information & Knowledge Management, 24*(3), 2530001.

Vos, L., & Armstrong, K. (2019). Context and process challenges associated with supervising postgraduate dissertations: An example from marketing. *International Journal of Management Education, 17*(1), 47–61. https://doi.org/10.1016/j.ijme.2018.11.005

Wright, A. (2024). Postgraduate supervision in a ChatGPT world: What's next? *International Conference on Higher Education Advances*, 480–487. https://doi.org/10.4995/HEAd24.2024.17244

Yagyaeva, E., Turobova, M., Nazirova, R., Daminova, Y., Merganova, N., & Oripova, N. (2024). A layout of bridging learning process to fill the gaps through AI to make the technical education for better access. *2024 4th International Conference on Advance Computing and Innovative Technologies in Engineering, ICACITE 2024*, 612–616. https://doi.org/10.1109/ICACITE60783.2024.10617424

CHAPTER TWO

# Generative AI as a 'Precipitant' of Challenges in Doctoral Supervision: A Dialogue Among Supervisors

Dimitar Angelov[1]
Luca Morini[2]
Israel Kariyana[3]
Nompilo Tshuma[4]

AFFILIATIONS

[1&2]Centre for Global Learning (GLEA), Coventry University, United Kingdom.
[3]Continuing Professional Teacher Development, Faculty of Education, Walter Sisulu University, Mthatha, South Africa.
[4]Centre for Higher and Adult Education, Stellenbosch University, Stellenbosch, South Africa.

CORRESPONDENCE

Email: dimitar.angelov@coventry.ac.uk

REFERENCE

Angelov, D., Morini, L., Kariyana, I., & Tshuma, N. (2026). Generative AI as a 'precipitant' of challenges in doctoral supervision: A dialogue among supervisors. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 17-37). ERRCD Forum. https://doi.org/10.38140/obp4-2026-02



**Abstract:** The role of supervisors has become more problematic since the rise of generative Artificial Intelligence (GenAI) in higher education and the disruption it has caused to teaching and research practices. This contribution employs a collaborative autoethnographic approach to articulate the various ways in which AI can act as a "precipitant" of challenges intrinsic to the position of PhD supervision in contemporary global academia, from the perspectives of four experienced supervisors based in the UK and South Africa. Drawing on their own experiences as PhD students, researchers, and supervisors of doctoral candidates, the authors shared possible answers and tensions using agreed-upon prompts to generate discussions that included the principles of academic integrity and ethical research practices that supervisors aim to cultivate in PhD students in the age of GenAI. Potential factors mitigating the misuse of the technology were discussed as part of this autoethnographic exploration, which coincidentally highlights the relational and chronological dimensions of the PhD, rather than the current hegemonic focus on "rushing to the end" and the final output.

***Keywords:*** Higher education, doctoral supervisors, doctoral students, GenAI, precipitant.

## 1. Introduction

With doctoral candidates occupying a liminal space between being students and being researchers (Keefer, 2015), the role of PhD supervisors can be seen as shaped by the competing expectations of guiding candidates through the institutionally driven process of earning a degree and facilitating their socialisation into discipline-specific discourse communities (Benmore, 2016). This complex task has become even more problematic since the rise of generative Artificial Intelligence (GenAI) in higher education and the disruption it has caused to teaching and research practices (Sardana, Fagan & Wright, 2023). As both 'teachers' and 'peers', supervisors need to assist their supervisees in navigating a landscape of diminished integrity and meaning-making rules, where new practices are yet to emerge and become established as normative.

There are certainly aspects of using GenAI as an assistive technology that hold promise for optimising doctoral supervision processes and outcomes. Dai et al. (2023, p. 84) found that "student independence [was] enabled by ChatGPT [and] allowed for a more proactive and autonomous approach to postgraduate research, where students were to take ownership of their research, gaining confidence and motivation." Dai et al. further conclude that the "ChatGPT-assisted research process encouraged the transformation of students from being passive apprentices to autonomous researchers" (2023, p. 84), a position that aligns with the recommendations of some leading publishers who advocate for the partial integration of GenAI into research writing (see, for example, Taylor & Francis, 2025).

At the same time, the increased use of GenAI tools for doctoral writing has coincided with a rise in academic misconduct cases (Cowling et al., 2023), reinforcing Badenhorst and Guerin's (2016) conclusion that writing and critical thinking are complex, nuanced, and interdependent processes. Furthermore, with regard to breaches of academic ethics and integrity, there are unresolved contradictions embedded in the very design of the technology itself: GenAI-assisted tools are trained on a vast amount of publicly available data harvested from the Internet (UNESCO, 2023, pp. 9–11) and recycled without proper attribution or a traceable pathway that can help identify the provenance of individual information items (Bowman, 2023). This sets an uncomfortable precedent for how others' words and ideas should be approached and integrated into research writing, making it difficult for supervisors to model good practice regarding academic-integrity compliance. In response to the ethical challenges intrinsic to GenAI tools and their outputs, Cowling et al. (2023) recommend that policy safeguards be established to ensure researchers can mitigate the lack of context and ethical framework, data bias, and equity concerns. Similarly, Wright (2024) urges greater clarity to address heightened weariness among academic staff who struggle with the wide-ranging implications of GenAI tools for their practice, in the absence of clear strategies for how to embrace or ban the technology.

The emphasis on increased productivity, which GenAI use seems to encourage (Watermeyer et al., 2024), aligns well with the need to augment the number of postgraduate qualifications in certain countries, such as South Africa, and to improve the quality of supervision practices (Govender and Markus, 2025). However, there is a danger that productivity will be framed in purely reductive terms as "generating outputs" (Becker & Lukka, 2023) at the expense of intellectual maturity and growth. Accelerated output generation is likely to compromise quality, which often comes from slow research (Kuus, 2015; O'Neill et al., 2014) and is essential for cultivating graduate attributes (Manathunga, Pitt & Critchley, 2009) that are and should remain at the heart of doctoral education.

In light of the above discussion, it follows that the literature remains ambiguous about the value of GenAI as a reliable assistive technology that can foster holistic postgraduate research development. We use this conclusion as a point of departure for our discussion in the present chapter, which intends to unpack some real and pressing concerns associated with the use of GenAI in doctoral education. To achieve this, we will employ a collective autoethnographic approach to articulate the various ways in which AI can act as a "precipitant" of challenges that are integral to the PhD journey in contemporary global academia, from the perspective of four experienced supervisors based in the UK and South Africa. For us, the technology is not solely a source of new issues, as it is often framed, but rather a factor that highlights existing structural dysfunctions in research and supervision.

Drawing on our own experiences of pursuing PhDs, working as researchers, and supervising doctoral candidates, we will share and comment on critical incidents that have shaped our perceptions of GenAI use in the context of doctoral education. Building on this, we will articulate answers and tensions related to prompts collectively agreed upon by all the authors. Through our choice of approach, we have aimed to avoid a simplified consensus on multifaceted problems and have instead constructed a dialogic narrative which, we hope, will serve as a catalyst for further academic conversations and the sharing of best practices.

### 1.1 Research questions

The study had three research questions:

- What are the supervisors' critical incidents involving AI that illustrate problems relevant to PhD supervision practice?
- What are the principles of ethical research that supervisors would want to cultivate in PhD students in the age of GenAI?
- What is the role of supervisors in developing PhD students' research skills in the era of GenAI, and whose responsibility is it to educate PGRs about the affordances of the technology?

## 2. Methodology

The authors of this article, who knew each other through previous contact, were brought together by their similar professional pathways as educators and doctoral supervisors, as well as a shared interest and concern about the role of GenAI as a disruptive technology, reconfiguring the landscape of research and doctoral supervision in their respective countries and globally. When commencing this study, we agreed to meet and explore the impact of GenAI on our scholarly and teaching practices, hoping to make a better, collective sense of the latest "technological creep" (Anderson, 2024) and the often subtle yet potentially fundamental effects it has on knowledge production and ownership. We also believed that by formally recording and disseminating our dialogue, we could offer prompts for further reflection to other colleagues in academia who are grappling with similar challenges and dilemmas.

The decision to engage in meaning-making conversations as a group led to the choice of methodology—collaborative autoethnography (CAE) (Chang et al., 2013; Chang, 2022)—which provided a suitable framework to capture all elements we deemed significant for our process. CAE is a form of qualitative enquiry in which the researchers are the participants in the study; it aims to capture individual experiences while calibrating these against the experiences of other researcher-participants and the existing literature on the problems under discussion (Chang, 2022; Lapadat, 2017). It allows for individual voices to emerge and be preserved as offering unique perspectives while forming a joint narrative that represents a collective experience and shared knowledge.

Our study is situated within the social constructionist paradigm, which privileges the human viewpoint in its framing of reality and conceives of knowledge as a dynamic process of collective meaning-making, whereby individuals interact within distinct sociocultural contexts (Crotty, 1998, pp. 52–57). Our collaborative autoethnographic method exemplifies the ontology and epistemology of social constructionism (Chang, 2013; Chang, 2022) and aligns well with our initial choice of topic, philosophical beliefs, and preferred mode of engagement as a group of researchers. CAE is a variant of autoethnography (AE) (Lapadat, 2017) or, more precisely, individual autoethnography (IAE) (Chang, 2013; Chang, 2022), and therefore conforms to the fundamental principles of the latter, such as the importance of lived experience as data and the choice of reflection and self-examination as analytical strategies (Adams et al., 2022). As CAE involves two or more participants/researchers, it offers an advantage over IAE by allowing those involved to access a wider evidence base for their conclusions and gain richer collective insights into the relevant issues (Lapadat, 2017).

With its multivocal nature (Lapadat, 2017), CAE is an apt choice for examining layered global phenomena, such as the impact of GenAI on doctoral education, as it accommodates a plurality of critical perspectives in a non-hierarchical manner. Among the unique strengths of this approach is that it facilitates a shared understanding and the co-creation of a joint narrative while preserving the distinctness and authenticity of individual voices and identities. As we are four doctoral supervisors and researchers from diverse backgrounds, institutions, and geopolitical settings, we felt it was important to find a framework that allows us to come together in our differences, which should be recognisable in our research output rather than hidden behind a uniform authorial persona, a standardised argument structure, and tone.

Our team was formed in response to a call for contributions to an edited collection, which was shared between two of the co-authors who decided on a preliminary topic for the chapter. This was then shared with the other two researchers, and a collaboration was forged to refine the thematic focus and submit an abstract to the editors of the collection.

Once the abstract was accepted, the four co-authors met to discuss and agree on a process for conducting the research and producing the written output. A joint decision was made that the study would comprise several online and hybrid discussions structured around collaboratively formulated prompts, reflecting the co-authors' shared experience and interests in the topic. In line with this plan, the co-authors met via MS Teams over nearly three months, first to negotiate and subsequently to explore the prompts through dialogue. The meetings were videorecorded and transcribed, using the MS Teams functionalities, to document the collaborative meaning-making and co-construction of knowledge that occurred as part of the study. The transcriptions were circulated among the co-authors, who used them as a basis for jointly writing the text of this chapter.

## 3.Findings and Discussion

The chapter's findings are presented and discussed under three prompts, following the research questions agreed upon during the early stages of the study.

### 4.1 Prompt 1: Critical Incident involving AI that illustrates problems relevant to PhD supervision practice

**Luca:** Ok, since I have a very recent and clear example, allow me to go first: we were doing the workshop with you, Dimitar, and I was interrogating ChatGPT, going quite deep into my specialism. It was trying for some "deep cuts" in the literature and history of the discipline, and the machine was keeping up quite well, though in the form of annoying bullet points. But there was still something that felt a bit off to me, and I couldn't really put my finger on it. Then we were discussing unfair advantage, and I asked it to draw an image of a very successful researcher, effectively cheating or fabricating research using AI to publish a lot more than others (as discussed by Majovsky et al., 2023). I had to ask for a satirical take to overcome its initial refusal, and it showed me an image of an old white-bearded man in a lab coat. And that's when I got it: that's how this machine writes, like the stereotype of the academic. I have nothing against this person, and I might be another white European man, but I don't write like that person; I know I can contribute something different. And my PGRs DEFINITELY don't look like that; they come from completely different experiences.

So, for me, that was a clear image, a blunt representation of epistemic injustice (see Fricker, 2017, for a discussion of this evolving concept). AI can provide a lot of information, but it can't take a stand. It always goes for "neutrality", which means it upholds the status quo. It's a bit of a ray of hope, really: while we need to acknowledge the hegemony of that stereotype in terms of academic writing, there's a lot more richness, a lot more diversity of voice that we can contribute and cultivate in PGRs.

**Nompilo:** Thank you for sharing what you took out of it, Luca. It's helped me rethink what I actually wanted to say. So my incident is actually from earlier this week. I'm writing a paper and was looking for a theory to help me frame the data I'm presenting. I tried searching in my existing theory toolkit, but couldn't find something that was a good fit. So, I started looking for a theory that would bring integrity and agency together.

I struggled quite a bit, but eventually found one – a newer theorist who is still writing to expand her theory. It turned out it had been a while since I read into a new theory for the first time. And you know how our theories are – and our theorists, usually abstract and often difficult to engage with the first time around (Jones, Bradbury & Le Boutillier, 2011). So I was struggling to understand it, but I could tell that this could be a really good fit for the data. Because I was struggling to grasp the theory's main tenets, I thought to myself, let me check my friends' [GenAI tools] and see what they say. Maybe they can explain it to me.

So I went to four of my friends and asked - so dudes this-and-this theory, help! The responses varied from simplified explanations to short, sometimes unintelligible summaries. Some, however, were a little helpful in clearing up the fog. I closed my laptop in a bit of a mental slump, questioning my philosophical intelligence. The next morning, I went back to the encyclopaedias that had introduced the theory to me, and it started clicking better, but maybe because I'd also given my brain time to process it. I then realised that my friend Chat had lied to me – slight inconsistencies that add up to a falsehood.

My biggest takeaway from this critical incident was understanding how students get into a theory by first of all figuring out what theory they need to use for their study, and then getting to a point where they understand it well enough to use it. Some challenges that students experience with learning to use social theory are linked to a mixed and sometimes conflicting explanation of what theory actually is or isn't – and therefore what comprises effective use in research (Abend, 2008; Hew et al., 2019; Sutton & Staw, 1995). Also, I experienced how GenAI tools could/could not assist in that regard – and the disciplinary expertise someone needs in order to ask useful prompts and critically interact with GenAI outputs.

**Dimitar:** Similarly to Nompilo, the critical incident I wish to reflect on arises from my own research practice – specifically, my experience of writing for publication. I had submitted a co-authored research article to a top-ranking international journal in education, and, after a long wait, I received the standard two peer-review responses. Reviewer One was succinct and extremely positive, requesting only a few minor, mostly technical changes to the text, while the proverbial Reviewer Two had produced copious feedback and a long list of major revisions. Upon closer inspection, the second reviewer's response seemed unnecessarily verbose and tautological, which only added to the disappointment and frustration any author would experience when having to rework their research output. Feelings aside, my collaborators and I worked very hard to address the reviewers' criticisms and resubmitted a rewritten article to the journal.

Then, as I was preparing to deliver a workshop on GenAI uses for research, the same workshop that Luca mentioned in his response, I realised that ChatGPT can produce a critique of any piece of writing that a user uploads, including a research paper. So, I tested this feature by using a randomly selected journal article, and the output I received looked very similar to the feedback of Reviewer Two, which was still fresh in my mind. Both texts had a similar layout, which included numbered headings highlighting standard points of an article critique, e.g. originality, focus, structure, etc., and sounded repetitive and verbose. Furthermore, the reviewer's feedback contained a peculiar orthographic symbol, an overly long dash, which is impossible to produce with a standard computer keyboard and which I had only seen in ChatGPT outputs. For illustration, I am copy-pasting the ChatGPT dash from another output, "challenging—even," which you can compare with a standard dash "–" and a hyphen "-" (all in bold). As you can see, the ChatGPT dash is longer and almost touches the words it links, a feature which sets it apart from the other two similar symbols.

Now, I know this is all circumstantial evidence, but I have no doubt in my mind that Reviewer Two had used GenAI to write, at least in part, his feedback. I don't think that the entire review was GenAI-produced, as it contained well-targeted recommended readings, but the very fact that some of it was shows the ubiquity of the technology in research practice. To be clear, the publishers of the journal, Taylor and Francis, permit limited use of GenAI in reviews – "to assist with improving review language" (2025) – however, a troubling thought persists that the reviewer had gone further than that. And even if they haven't, given the repetitiveness of the review, GenAI had not been taken full advantage of.

I thought it was important to share this experience because it reflects well the research landscape our PhD students will have to navigate henceforth: a brave new world where GenAI will provoke fresh challenges and anxieties, but also offer opportunities for different, perhaps better, ways of doing research.

**Israel:** I have mixed encounters with GenAI, which are not necessarily critical incidents. My first encounter was after encouragement from a good colleague in another faculty to explore the potential benefits of ChatGPT. That came after successive discussions about the emerging trends in education and, generally, in life, and the associated concerns. The 'tech-savvy' colleague even presented scenarios in which GenAI was abused in the medical and law professions, and the costly repercussions that followed. However, when he illustrated the way he uses ChatGPT almost daily, and the manner in which I could benefit from utilising it, that made me ponder whether I was neither too rigid nor too difficult for myself. Thereafter, my overall experience, drawn especially from supervision, is that GenAI's responses are quite generic; they demand contextualisation. This confirms Wright's (2024) stance that, adding in nouveau complexities of contract cheating, Artificial Intelligence using ChatGPT is a game changer and disruptor in HE; academic weariness is heightened as many universities try to better understand AI and ChatGPT. The vastness and implications of these AI tools are not clarified for most staff, with few policies or clear strategies in place to work with it, against it, embrace it, or ban it. Interestingly, Dai, Lai, Lim, and Liu (2023) found that Australian postgraduate students believed ChatGPT enhanced the quality and pace of their work. As such, my approach to using GenAI is two-faced: reluctant but motivated by the desire to stay 'informed' and ahead of, especially my students, and I believe that has benefited me proportionally.

However, can I also briefly comment on your critical incidents? So, for both of you, Luca and Nompilo, you received outcomes that were/are not reflective of your ontological and epistemological orientations. But then, I am not sure whether such outcomes have anything to do with the quality of the questions prompted, which, of course, relate to the follow-up questions. The reality is that asking different questions may yield different outcomes, but again, as I said earlier, my experience has been that most outcomes are not very relevant when the responses required come from unfamiliar cases. Cowling et al. (2023) recommend that policy safeguards are needed to ensure research responses address a lack of context, data bias, equity concerns, and a lack of an ethical framework. In my view, however, the good friend GenAI relies on the quality of data available at any given time, and this data is incrementally being fed into the system at an alarming rate due to technological advancements and what looks like a contagious global interaction with GenAI. This is in sync with Wright's (2024) orientation that HE institutions are currently challenged to provide guidelines/best practice for supervisors because of the advancement in the speed of machine learning. By the time a document is written and released to staff, AI has galloped on, leaving academics in its wake.

### 4.2 Prompt 2: Modelling ethical research practice (authorship, ownership of ideas, originality, contribution)

**Nompilo**: Speaking to the point regarding our interactions with GenAI tools (or other human resources), I've had long discussions with my doctoral students about this. If you look at the fact that the thesis itself has a reference list, we know those are not the only books or articles that the student has read. The student has had to read much, much more, maybe double or triple, or even five times more than what they actually include there, which has informed their thinking and has influenced the direction they've taken in their research. Sometimes you find a bibliographic list, but even then that's not a complete list. So what does that say about what we can and can't reference? Should GenAI tools be included?

Another issue is the fact that our institution now has guidelines on what you need to do if you use any GenAI tool in your research, in your writing assignments, and so on. There's a little table that you complete about what you used, what you used it for, and what that means for the authorship of the document. Is it still yours? How is it still yours? And I'm thinking of things like Grammarly and NVivo, and language editors and so on – all these AI, technical, and human resources have helped me make my work better in some way. But I've hardly seen mention of their use in published or other research work. So how do we decide what to include and what not to include, or who to include and who not to include?

The writing process is complex, and the thinking behind it and what adds to that thinking is also quite complex and very nuanced (Badenhorst & Guerin, 2016). So we're asking questions without answers, I think. And there's clearly still much more we need to think about as universities and as supervisors.

**Dimitar**: Supervisors are perceived as mentors in good research practice, including ethics and integrity, and should ensure that PhD students comply with institutional and disciplinary norms. Although seemingly straightforward, the reality of observing ethics and integrity rules can often be subjective and ambiguous. A concise yet telling example is the often discretionary nature of knowledge attribution in academic writing. As an operational principle, common knowledge within a discipline does not normally require attribution; however, there is no clear consensus on which knowledge is owned collectively and which individually (Shi, 2011).

I believe that the contested nature of ownership and, by extension, authorship of ideas, about which we, as supervisors, need to educate our PhD students, is compounded further by the adoption of GenAI for research. It is a well-known fact that AI-assisted tools are trained on vast amounts of publicly available data harvested from the Internet (UNESCO, 2023, 9-11), which is recycled without proper attribution or even a traceable pathway that can help identify the provenance of individual information items (Bowman, 2023). The design of the technology itself is thus a major disruptor of our understanding of shared or individually authored knowledge and what constitutes ethical behaviour in navigating intertextuality in academic meaning-making.

With GenAI being adopted as a multi-purpose research tool, our long-standing principles of authorship are being redefined or rendered obsolete in multiple other ways. One among many examples is the use of the technology to explore ideas or generate plans, outlines, or early drafts, which then feed into the researcher's textual output. This practice is both already established amongst PhD students (English et al., 2025) and recommended as pedagogically sound by UK education advisory bodies such as the Quality Assurance Agency (QAA) (May 2023) and JISC (2023). For me, the process of idea generation and development with GenAI is problematic in terms of our current understanding of authorship, as it constitutes an active intervention into the creative aspect of research (see also Angelov, 2025 – forthcoming). Incorporating pre-fabricated argument segments into your writing, even if retrospectively referenced, is similar to co-authorship which remains unacknowledged. Perhaps, in the future, such co-creation between human intelligence and AI will become the norm, as Eaton's idea of "postplagiarism" suggests (2024); however, this would require a redefinition not just of the concept of authorship but the entire knowledge economy which is underpinned by legal categories based on identifiable provenance and ownership of ideas, such as copyright and intellectual property.

**Luca**: Thanks Dimitar, and I want to say that I completely take your point about the issue of ownership. I am not the biggest fan of intellectual property, but I am a stickler for genealogy (in the Foucauldian sense) - so while I would like a complete shift in the political economy of academic publishing, where GenAI is really a strong precipitant for disruption, I would still want to know who wrote something, in what socio-historical context, and from what standpoint. In connection with this, the modelling point for me is very interesting, because of the approach to supervision that I saw here in the UK and in South Africa, vis-à-vis what I experienced in Italy (for a more in-depth discussion of different PhD models, see Dominguez-Whitehead & Maringe (2020)). Here, I haven't worked with PGRs on projects or on teaching: full-time PGRs spend most of their time on their individual projects, and part-timers barely have the time to do their project and their daily job, so I don't have many opportunities to "model", to constitute that "genealogy" that I mentioned above in terms of the evolution of thought. It's something which I am honestly a bit sad about, not having that ongoing collaboration and conversation.

But for me, it loops back again to enabling different styles of writing, looking different from what is established and from what I myself write - to push the boundaries of the thesis in terms of both content and style. I was thinking of Fanon's "Black Skin, White Masks", which might well be in the top 10 most influential books of 20th-century social science, and was rejected as a thesis.

There was (and there still is much) structural racism, but it's about what is considered acceptable in terms of new knowledge. I don't think these tools can help us push these boundaries in that same way - by definition, they can only replicate. And we know that undergraduates, PGRs, supervisors (and even reviewers, as we have seen!) use them, because we all have deadlines, and we see that we can rely on these tools to get through them, because we are all tangled up in "generating outputs" and productivity demands (as discussed by Becker & Lukka, 2023) that only tangentially have to do with advancing knowledge, with continuing the genealogy that I was referring to.

The more I think that these tools could help us with the core drive of our work, the more I feel that this core drive has been warped. If one can use these tools to be what counts as a successful scholar nowadays, then I have serious problems with the notion of a successful scholar - again, there's an ontological break with the genealogy of thought. And of course, I need to put food on the table, so I have to be that person to a degree. But I still can say, "look, maybe this is not the right direction".

**Israel**: To my students, I'm saying, do you know the fundamentals? Do you know if I would ask you to do this without this 'assistant'? Would you get it right? You know you can't continue to develop students the 'traditional' way if they have other ways. Supervisors often engage with their students based on their own prior experiences as postgraduate students (Govender & Markus, 2025). I'm supervising many students. So, as they present me with some work, I'm asking, is this their work, like I'm asking? And when I'm asking them a few clarity-seeking questions to establish whether they can respond as anticipated, at least to my satisfaction, based on the stage and level of study, some of them are getting challenged. So, I'm then also querying how they arrived at presenting to me something that they struggle to at least articulate. I have realised that ChatGPT has a particular way of expressing itself, coupled with its out-of-context responses, a pattern which I established in some of my students' writings. I mean, I don't know for sure what the good friend gives them, but for some of them, it is coming out that there is no coherence in their submissions. This experience contradicts findings by Dai et al. (2023, p. 84) that "student independence enabled by ChatGPT allowed for a more proactive and autonomous approach to postgraduate research, where students were to take ownership of their research, gaining confidence and motivation. Moreover, by addressing technical tasks independently, students had more time to engage in deeper critical thinking and plan for the bigger picture of their research project. The ChatGPT-assisted research process encouraged the transformation of students from being passive apprentices to autonomous researchers." So, I'm always engaging some of them, saying I'm not hearing your voice, or you don't seem to be owning up to your work. This aligns with Govender and Markus (2025) who argue that the pressure to increase the number of postgraduate qualifications in South Africa has intensified the need for quality supervision within universities.

That has led me to conclude that they are going all the way to use this good friend, but then they are not going back to do what they're supposed to do. Therefore, this is the centre of some of our ongoing group conversations with the students I supervise, with one strong recommendation of the need to set self-limits on choices of how to benefit from using GenAI, with a key priority to remain as the authors and owners of their individual work. That is, the human part must remain human, with evidence of its existence in the work, and understanding that machines will continue to be machines, just putting it bluntly.

### 4.3 Prompt 3: Supervisors' role in developing research skills (how to write with GenAI; who has the expertise and should be responsible for educating PGRs about the affordances of the technology)

**Israel**: The question is how can/do we as supervisors take our students forward during these unprecedented times? Because sometimes as supervisors, we are also, you know, sceptical. Yet, sometimes these students are way ahead and they may have all this information. So at the end of the day, I ponder to say, do we have to limit them, and how? If not, then, (how) do we support them to ethically use this AI? This, to me, is the maze in which supervisors find themselves enclaved by GenAI. So, we are saying this AI has precipitated the challenges that supervisors are facing. Wright (2024:480) also raised the same, asking, "With the advent of ample search engine opportunities, GenAI technology, and adhering to proper academic integrity processes, how can supervisors navigate these complexities, unsupported by solid procedures, due to the rapidly changing nature of AI?"

In recent years, there have been considerable examples of substantial cheating in doctoral studies (Cowling et al., 2023). Such conversions need to continue to be discussed in academia with the understanding that knowledge expansion and advancement must be real. AI enables a student ethically cleared to unethically complete a thesis/dissertation, from chapter one development to the final chapter, in less than three months. Or to start and complete a systematic literature review paper in a couple of days or so. But, is that authentic knowledge production? Even after critiquing the work, sometimes this problem of artificial knowledge production escapes the tentacles of supervisors or reviewers. Hence, I continue to emphasise the need to respect the academic space by understanding that at the doctoral level, or any other level for that matter, contribution to knowledge and defending one's space within a discipline illuminate the ultimate definition of a doctor. In one study, Caillaud and Skec (2024) found that a subset of interviewees cautioned against treating GenAI as a mere "search engine" and expressed their doubts regarding the quality of output for literature review, attributing their concerns to the quality of the sources employed in training the GenAI and the potential for "hallucinatory" content.

So, sometimes it is important to close out technology and go back to (literally) doing it manually, because I want to demonstrate to them how they have to get from the first step. I'm saying, do you know the fundamentals? I keep on thinking about what the students could do without AI, and what they could do with AI. That is the launchpad of my next questions to my students: How much do you know if I were to ask you to do this without the assistance of this machinery? Can you explain how you would approach it, and when you get what you get, would you be able to critique it? And how? Would you be able to establish whether it is correct or wrong? Chauke, Mkhize, Methi and Dlamini (2024)'s study recommends the immediate development of an innovative policy on the ethical use of AI at South Africa's historically disadvantaged universities. This policy should emphasise ethical guidelines for postgraduate students when utilising AI tools such as ChatGPT to ensure responsible and effective integration into their academic success.

However, my view is that the policy ought to apply across South Africa's higher education sector. In line with such a position, just to be careful whenever my students present some work to me, I'm always asking myself if it is their work or another victimisation by AI, and then, let me use the simplest way to be convinced that they own the work. Cowling et al. (2023) posit that although expectations are significant on the ethical conduct of research students, this is not always the case, and effective detection will be needed to respond to a minority of students using the AI tool incorrectly. So, from my point of view, I am consistently indicating that we should continue asking ourselves: how far can we go without AI, and how far can we go with AI? Caillaud and Skec (2024) suggest that, rather than seeking to prevent PhD students from utilising such tools, supervisors should equip them with the knowledge to utilise these new GenAI resources effectively and ethically. However, as a prerequisite, supervisors themselves must become familiar with the diverse functionalities offered by GenAI research tools and attain proficiency in their application. For me, this seems to suggest that to be helpful in assisting with GenAI, supervisors have to be a step ahead of the students. But is this not an extra layer of supervisory responsibilities, and how much more can supervisors contribute to developing students' acquaintance with AI, as we also have our own challenges, and we also have to draw the line as supervisors?

**Luca**: I like your questioning, Israel, and in fact, I think that a lot of the issues you raise would be less urgent if academic cultures (and the metrics we sadly have to live by) weren't so writing-centric and more about real conversations in real contexts. Maybe it's an excessive position, and I am going to sound like Socrates, when (as narrated in Plato's *Phaedrus*) he argued that writing is going to ruin knowledge. Maybe I am already an old man rambling about what kids do nowadays, but I struggle to see the link between GenAI and research skills, or "doctorateness," whatever we take that to mean (and I agree with Wellington (2013) that we always need an ongoing debate on that matter), and therefore I don't think it falls within my responsibilities as a supervisor to teach them about the technology, if not in negative terms.

This is because I just don't see how these tools can help the PGRs get to "doctorateness," which I personally see as the professional capacity to make value judgments—effectively, to take a justified, rigorous public stand, which of course includes a matter of voice and positionality, something that I touched upon during my critical incident (in doing this, I subscribe to a version of Hekman's (1997) standpoint theory).

I mean, of course AI can help in technical ways, in the same way that we are using a technological tool now to have this conversation at a distance. And of course the choice and affordances of the platform shape it in some way (Vidolov, 2022), but it is not in itself the conversation, the knowledge-generating dialogue itself. Or an even worse example: I can use a microwave to quickly heat lunch, and that will save me time to focus on my PhD; it will help me towards it in a real sense, but I wouldn't say that the microwave per se helps me get to "doctorateness," and therefore that I should teach PGRs how to use a microwave.

So to summarise, I don't feel like these tools can help, if not with some of the logistics, and I don't feel that it necessarily falls upon me, as a supervisor, to take care of that (though of course I am happy to help and share what I do, for example, in terms of managing references).

**Nompilo**: So, prior to November 2022, I looked at the things we'd talked about in our postgraduate supervision workshops. We talked about things like academic writing, an ethical mindset, criticality and a whole range of skills and knowledge – the graduate attributes (Manathunga, Pitt & Critchley, 2009). These attributes comprise the identity that doctoral students should have by the end of the degree, and I still think that is the responsibility of the supervisor – enabling the development of those attributes. But now there's the aspect of GenAI and the tools in that space. Already for me there's a tension there. Most of the tools are not actually supporting the development of any of those graduate attributes at doctoral level. They are supporting the development of a particular output. However, the output itself needs to be a demonstration of the attainment of the graduate attributes.

So for me, I see the challenge stemming from our almost exclusive focus on the product – rather than the process. As soon as students come in, we already start demanding certain outputs from them which will start to demonstrate that they are developing doctorateness or the doctoral identity (Trafford & Leshem, 2009). In the workshops I run with postgraduate students, I try to avoid demonstrating GenAI tools and rather focus on the principles of ethical research – and the possible 'arrested development' depending on how they use GenAI to support the production of outputs that we demand of them.

**Dimitar**: It is great that we have different takes on this prompt, which is revealing of our individual approaches to supervision, but also of the intrinsic ambiguities of the supervisor-supervisee relationship. I share Luca's scepticism when it comes to the use of GenAI to bolster the neoliberal discourse of productivity in higher education, a problem that has already been identified by scholars interested in the research uses of the technology (Watermeyer et al., 2024), and I agree there is a real danger that accelerated output generation will compromise quality, which often comes with slow research (Kuus, 2015; O'Neill et al., 2014). However, if GenAI is here to stay, it will inevitably transform the process of research and writing, which will require the development of new skill sets geared towards the production of co-created research outputs. Regardless of the exact practical and ethical implications of this transformation, I believe the successful researcher of tomorrow will need a different type of digital literacy that combines advanced technical and critical thinking skills.

Going back to the original prompt, I find it difficult to separate research skills from the intellectual work that goes into a PhD thesis, as the latter is very much predicated on the former. In fact, I believe that supervisors always teach skills, albeit implicitly, when they give feedback on formal aspects of students' writing, e.g. argument organisation, sentence and paragraph structure, or style, which I am sure we all do. The key distinction to bear in mind here is between implicit and explicit teaching of these skills, which, again, is a grey area in the remit of the supervisor's role. How much explicit support for research skills development is fair to expect from a supervisor who is first and foremost a subject expert and may not have relevant pedagogical or technical expertise?

ChatGPT is bringing this problem into sharper focus for me as it works as an amplifier of students' skills-development needs. If the technology is a game changer and its expert use will determine the professional success of future researchers, shouldn't we all be teaching our students how to unlock its full potential? Personally, I feel I will be short-changing my students if I don't do so. Furthermore, and this goes back to doctoral studies being degree programmes, we need to bear in mind that the PhD thesis is, after all, an elaborate piece of coursework which is submitted for assessment. In recent years, there has been a concerted effort as part of the Scholarship of Teaching and Learning to educate students about the assessment tasks they are required to complete; in other words, to develop students' "assessment literacy" (Price et al. 2012). I find the concept of "assessment literacy" to be extremely helpful as it draws attention to the often tacit knowledge, skills, and competencies involved in performing an assessment task, which are either absent from consideration or neglected in favour of subject knowledge. If a fair approach to teaching involves educating students about these technical and procedural elements of assessment, then teaching them how to use GenAI for research should be an integral part of doctoral training and supervision.

## 4. Conclusion and Recommendations

When teasing out common threads in our critical incidents and in how we model academic roles and support the development of doctoral capabilities, our discussions have consistently highlighted pre-existing points of tension and grey areas between different stakeholders: the growing, output-oriented productivity emphasis in academia (Niyazova, 2021); the lack of formalisation in the role of the supervisor; and the complexities of modelling an evolving role for PGRs (Guarimata-Salinas, Carvajal & Jimenez-Lopez, 2024); and the inequality of ownership and authorship in publishing (Hart & Perlis, 2021).

While all of these predate GenAI, our conversations have also revealed how GenAI technology is quickly bringing these pre-existing issues into much sharper contrast (Bozkurt, 2024). With the ongoing radical change in academia's meaning-making matrix, the question arises as to whether a breaking point is approaching. Although the narrative of "fast social change" has been rightfully subjected to strong critique and scrutiny, including for its reactionary overtones (Lawson, 2014; Kavanagh, Lightfoot & Lilley, 2021; Ballard & Barnett, 2021), our narratives seem to frame the advent of GenAI in academia as a one-off case of "a frog in boiling water" (Steffan, 2024), which has the potential to elicit radical critiques and reactions (McQuillan, 2022).

There is thus an opportunity for transformative potential in engaging in conversations about GenAI in HE: by identifying the technology as a "precipitant", we highlight both a historical continuity in connection with existing structural dysfunctions and a radical acceleration in how these are experienced and technologically mediated in our supervisory practices. In doing so, we can envisage ways of mitigating some of the most impactful effects, and hopefully, securing a stronger position to articulate and raise awareness of those structural issues in our supervisees.

Potential factors for mitigation were therefore discussed as part of this autoethnographic exploration, which highlight the relational and chronological dimensions of the PhD, rather than the current hegemonic focus on "rushing to the end" and the final output. In line with this, we can draw a few potential recommendations for supervisory practice from our conversations, all pertaining to "staying with the trouble" (Haraway, 2016): not neat solutions of reconciliation with technology or of recuperation of one-sided epistemic authority, but rather plural aspects of the ongoing, more modest, but more pervasive processes of "getting on together" (ibid.) in a troubled academic ecosystem. These are as follows:

- As Nompilo suggests, recovering a more process-oriented thinking around the PhD by limiting instrumental strategies that seem to funnel candidates into a long series of discrete outputs would weaken the "pull" that GenAI technology exerts. This approach echoes Manathunga et al.'s (2021) critique of the focus on efficiency and outputs, as well as their warping effect on doctoral temporalities.
- Emphasising once again the relational dimension of doctoral study, both Israel and Luca affirm the value of actual, back-and-forth verbal conversations and questioning with supervisees, and more broadly with the academic community and the general public. This

explicitly pertains to recovering the embodied character of knowledge production, which is historically, geographically, and culturally contextualised, and opposing a "gaze from nowhere" approach that is often insidiously inherent to algorithmic production (Haraway, 1988).

- Finally, and again in connection with enhancing the capacity of doctoral candidates to engage productively in academic discourse, Dimitar highlights the importance of cultivating assessment literacy, including how GenAI can (and cannot) address the formal requirements of the PhD. This will involve developing a dynamic awareness of the inevitable transformations that AI technologies will bring to how academic integrity and academic identities are framed and perceived (Hellaway, 2025).

Ultimately, there appears to be a potential way to positively address and mitigate the heightened challenges brought by GenAI in doctoral supervision: create relational spaces where we, as a community of supervisors and postgraduate researchers, acknowledge the thesis and "doctorateness" as an emergent property of conversations marked by shared understanding and co-creation of a joint narrative. This approach echoes African philosophies such as Ubuntu, which highlight collective knowledge creation(see Kgari-Masondo et al., 2024, for a specific discussion in the context of doctoral authorship), while preserving the distinctness and authenticity of individual voices and identities in their respective and distinct roles within this process.

## 5. Limitations and Future Research

In choosing to engage in CAE, we accepted the limitations that come with this methodology. Firstly, the number of participants in a collective autoethnographic study is limited compared to other qualitative methods, starting from as few as two (Lapadat, 2017; Chang, 2022). However, what might be perceived as restricted breadth is typically offset by the depth of the research data—collected over iterative conversations—and the agency that participants gain when actively participating in the co-creation of knowledge as co-authors.

Secondly, CAE makes it difficult to safeguard anonymity, and we made a conscious decision to disclose our identities when reporting on our individual experiences and beliefs. A similar approach was taken by many other collaborative autoethnographic studies (e.g., Guerrero-Nieto & Quintero-Polo, 2024; Hernandez et al., 2015; Sheridan et al., 2020). In doing so, we also aimed to communicate our cultural and gender positionality, which is an intrinsic part of our professional and life stories. We are aware that by foregoing anonymity, we have created a certain vulnerability for ourselves as researchers, but also for others who may be implicated in our narrative (Chang, 2022). To minimise any possibility of negative exposure, we have followed "the ethical principle of 'do no harm'" (Chang, 2022, p. 62) to anyone involved by withholding personal identifiers and potentially compromising information. In future studies, other researchers may want to use anonymised data to investigate high-risk topics, such as the intentional misuse of GenAI in both doctoral and research writing.

Finally, we acknowledge the disciplinary limitations of our scholarly and teaching practice. As social scientists specialised in education, we can speak for a particular academic community, but we recognise that colleagues from other disciplines may have different views and identify different challenges in relation to GenAI use in doctoral education. We hope our conversation provides insight and prompts others to critically reflect on the adoption of this technology in the meaning-making and knowledge-production practices of their disciplines.

## 6. Declarations

**Funding:** This research did not receive any external funding.
**Conflicts of Interest:** The author(s) declare no conflict of interest.
**Use of Artificial Intelligence:** The current work was created without the assistance of artificial intelligence technologies, as confirmed by the author(s).

## References

Abend, G. (2008). The meaning of 'theory'. *Sociological Theory, 26*(2), 173-199. https://doi.org/10.1111/j.1467-9558.2008.00324.x

Adams, T. E., Holman Jones, S., & Ellis, C. (2022). Introduction: Making sense and taking action: Creating a caring community of autoethnographers. In T.E. Adams, C. Ellis, & S. Holman Jones (Eds.), *Handbook of autoethnography* (pp. 1–19). Routledge. https://doi.org/10.4324/9780429431760-1

Anderson, M. (2024). *Public education in the digital age: Neoliberalism, edtech, and the future of our schools*. Routledge.

Angelov, D. (2025). Developing academic integrity-compliant regulations and policies on the use of generative AI in higher education: Insights from the United Kingdom. In A. E. King (Ed.), *Artificial intelligence, pedagogy and academic integrity* (pp. 129–141). Springer. https://doi.org/10.1007/978-3-031-92534-4_9

Badenhorst, C., & Guerin, C. (Eds.). (2016). *Research literacies and writing pedagogies for Master's and Doctoral writers.* Brill Academic Publishers.

Ballard, R., & Barnett, C. (2022). Apprehensions of social change. In *The Routledge Handbook of Social Change* (pp. 1–16). Routledge.

Becker, A., & Lukka, K. (2023). Instrumentalism and the publish-or-perish regime. *Critical Perspectives on Accounting, 94*, 102436.

Benmore, A. (2016). Boundary management in doctoral supervision: How supervisors negotiate roles and role transitions throughout the supervisory journey. *Studies in Higher Education, 41*(7), 1251–1264. https://doi.org/10.1080/03075079.2014.967203

Bowman, S. R. (2023). Eight things to know about large language models. *arXiv.* https://doi.org/10.48550/arXiv.2304.00612

Bozkurt, A. (2024). GenAI et al. Cocreation, authorship, ownership, academic ethics and integrity in a time of generative AI. *Open Praxis, 16*(1), 1-10. https://doi.org/10.55982/openpraxis.16.1.654

Caillaud, E., & Skec, S. (2024). Supervision of design PhD students in an era of generative artificial intelligence. In *26th Conference on Engineering and Product Design Education (E&PDE 2024), The Design Society; Institution of Engineering Designers, Sep 2024, Birmingham (UK), United Kingdom* (pp. 420-425). hal-04766967

Chang, H. (2013). Individual and collaborative autoethnography as method: A social scientist's perspective. In S. Holman Jones, T. E. Adams, & C. Ellis (Eds.), *Handbook of autoethnography* (1st ed., pp. 107-122). Left Coast Press.

Chang, H. (2022). Individual and collaborative autoethnography for social science research. In T.E. Adams, C. Ellis, & S. Holman Jones (Eds.), *Handbook of Autoethnography* (2nd ed., pp. 53–65). Routledge. https://doi.org/10.4324/9780429431760-6

Chang, H., Ngunjiri, F. W., & Hernandez, K.-A. C. (2013). *Collaborative autoethnography*. Routledge. https://doi.org/10.4324/9781315432137

Chauke, T., Mkhize, T., Methi, L., & Dlamini, N. (2024). Postgraduate students' perceptions of the benefits associated with artificial intelligence tools for academic success: In case of ChatGPT AI tool. *Journal of Curriculum Studies Research*, *6*(1), 44-59. https://doi.org/10.46303/jcsr.2024.4

Cowling, M., Crawford, J., Allen, K.-A., & Wehmeyer, M. (2023). Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research supervision. *Australasian Journal of Educational Technology*, *39*(4), 89-103. https://doi.org/10.14742/ajet.8598

Crotty, M. (1998). *The foundations of social research: Meaning and perspective in the research process* (1st ed.). Routledge. https://doi.org/10.4324/9781003115700

Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology*, *39*(4), 74-88. https://doi.org/10.14742/ajet.8843

Dominguez-Whitehead, Y., & Maringe, F. (2020). A cross-national analysis of PhD models. *International Journal of Comparative Education and Development*, *22*(3), 233-245. https://doi.org/10.1108/IJCED-01-2020-0003

Eaton, S. E. (2023). Postplagiarism: Transdisciplinary ethics and integrity in the age of artificial intelligence and neurotechnology. *International Journal for Educational Integrity, 19*(1), Article 23. https://doi.org/10.1007/s40979-023-00144-1

Ellaway, R. H. (2025). What do we become? Artificial intelligence and academic identity. *Medical Teacher*, *47*(10), 1569-1570. https://doi.org/10.1080/0142159X.2025.2523468 English, R., Nash, R., & Mackenzie, H. (2025). 'A rather stupid but always available brainstorming partner': Use and understanding of generative AI by UK postgraduate researchers. *Innovations in Education and Teaching International.* https://doi.org/10.1080/14703297.2024.2446236

Fanon, F. (2008). *Black skin, white masks* (R. Philcox, Trans.). Grove Press. (Original work published 1952)Fricker, M. (2017). Evolving concepts of epistemic injustice. In *The Routledge handbook of epistemic injustice* (pp. 53-60). Routledge.

Govender, N., & Markus, E. D. (2025). The bumpy road to postgraduate supervision in South Africa: A collaborative autoethnography. *International Journal of Educational Research Open, 9*, Article 100461. https://doi.org/10.1016/j.ijedro.2025.100461

Guarimata-Salinas, G., Carvajal, J. J., & Jiménez López, M. D. (2024). Redefining the role of doctoral supervisors: A multicultural examination of labels and functions in

contemporary doctoral education. *Higher Education, 88*(4), 1305–1330. https://doi.org/10.1007/s10734-023-01171-0
Guerrero-Nieto, C. H., & Quintero-Polo, A. H. (2024). Walking toward the demonumentalization of qualitative research: A collaborative autoethnography account while producing an educational podcast. *The Qualitative Report, 29*(2), 464–483. https://doi.org/10.46743/2160-3715/2024.6275
Haraway, D. (1988). Situated knowledges: The science question in feminism and the privilege of partial perspective. In *Feminist Studies, 14*(3), 575-599. Routledge. https://doi.org/10.2307/3178066
Haraway, D. J. (2016). *Staying with the trouble: Making kin in the Chthulucene*. Duke University Press. https://doi.org/10.1215/9780822373780.
Hart, K. L., & Perlis, R. H. (2021). Authorship inequality: A bibliometric study of the concentration of authorship among a diminishing number of individuals in high-impact medical journals, 2008–2019. *BMJ Open, 11*(1), e046002. https://doi.org/10.1136/bmjopen-2020-046002
Hekman, S. (1997). Truth and method: Feminist standpoint theory revisited. *Signs: Journal of Women in Culture and Society, 22*(2), 341-365.
Hernandez, K.-A. C., Ngunjiri, F. W., & Chang, H. (2015). Exploiting the margins in higher education: A collaborative autoethnography of three foreign-born female faculty of colour. *International Journal of Qualitative Studies in Education, 28*(5), 533–551. https://doi.org/10.1080/09518398.2014.933910
Hew, K.F., Lan, M., Tang, Y., Jia, C., & Lo, C.K. (2019). Where is the "theory" within the field of educational technology research? *British Journal of Educational Technology, 50*(3), 956-971. https://doi.org/10.1111/bjet.12770
Hyland, K. (2016). *Teaching and researching writing* (3rd ed.). Routledge.
Jisc. (2023, July 12). *Generative AI - a primer*. Retrieved July 5, 2024, from https://www.jisc.ac.uk/reports/generative-ai-a-primer
Jones, P., Bradbury, L., & Le Boutillier, A. (2011). *Introducing social theory* (2nd ed.). Polity Press.
Kavanagh, D., Lightfoot, G., & Lilley, S. (2021). Are we living in a time of particularly rapid social change? And how might we know? *Technological Forecasting and Social Change, 169*, 120856. https://doi.org/10.1016/j.techfore.2021.120856
Keefer, J. M. (2015). Experiencing doctoral liminality as a conceptual threshold and how supervisors can use it. *Innovations in Education and Teaching International, 52*(1), 17-28. https://doi.org/10.1080/14703297.2014.981839
Kgari-Masondo, M. C., Chingwanangwana, B., Chimbunde, P., Zondi, L. P., Umejei, A. L., & Mosina, D. F. (2024). Whose PhD is it? A case of supervision in higher education in South Africa. *E-Journal of Humanities Arts and Social Sciences*, *5*(14), 2597–2612. https://doi.org/10.38159/ehass.202451412
Kuus, M. (2015). For slow research. *International Journal of Urban and Regional Research*, *39*(4), 838–840. https://doi.org/10.1111/1468-2427.12257
Lapadat, J. C. (2017). Ethics in autoethnography and collaborative autoethnography. *Qualitative Inquiry*, *23*(8), 589–603. https://doi.org/10.1177/1077800417704462
Lawson, T. (2014). A speeding up of the rate of social change? Power, technology, resistance, globalisation and the good society. In *Late modernity: Trajectories towards morphogenic society*

(pp. 21–47). Springer International Publishing. https://doi.org/10.1007/978-3-319-03266-5_2

Májovský, M., Černý, M., Kasal, M., Komarc, M., & Netuka, D. (2023). Artificial intelligence can generate fraudulent but authentic-looking scientific medical articles: Pandora's box has been opened. *Journal of Medical Internet Research*, *25*, e46924. https://doi.org/10.2196/46924

Manathunga, C., Pitt, R., & Critchley, C. (2009). Graduate attribute development and employment outcomes: Tracking PhD graduates. *Assessment & Evaluation in Higher Education*, *34*(1), 91–103. https://doi.org/10.1080/02602930801955945

Manathunga, C., Qi, J., Bunda, T., & Singh, M. (2021). *Time mapping: Charting transcultural and First Nations histories and geographies in doctoral education.Discourse: Studies in the Cultural Politics of Education*, *42*(2), 215–233. https://doi.org/10.1080/01596306.2019.1603140

McQuillan, D. (2022). *Resisting AI: An anti-fascist approach to artificial intelligence.* Bristol University Press.

Niyazova, M. V. (2021). Individual academic productivity vs new managerialism in academic research. *University Management: Practice and Analysis*, *25*(2), 114–122. https://doi.org/10.15826/umpa.2021.02.018

O'Neill, M., Martell, L., Mendick, H., & Müller, R. (2014). Slow movement/slow university: Critical engagements: Introduction to the thematic section. *Forum, Qualitative Social Research, 15*(3), Art. 16. https://nbn-resolving.de/urn:nbn:de:0114-fqs1403166

Price, M., Rust, C., O'Donovan, B., Handley, K., & Bryant, R. (2012). *Assessment literacy: The foundation for improving student learning.* Oxford Brookes University.

Quality Assurance Agency (QAA). (2020). *Characteristics statement: Doctoral degree.* https://www.qaa.ac.uk/the-quality-code/characteristics-statements/characteristics-statement-doctoral-degrees

Quality Assurance Agency (QAA). (2023, May 8). *Maintaining quality and standards in the ChatGPT era: QAA advice on the opportunities and challenges posed by generative artificial intelligence.* https://www.qaa.ac.uk/en/membership/membership-areas-of-work/generative-artificial-intelligence/qaa-advice-and-resources

Sardana, D., Fagan, T. R., & Wright, J. T. (2023). ChatGPT: A disruptive innovation or disrupting innovation in academia? *The Journal of the American Dental Association, 154*(5), 361-364. https://doi.org/10.1016/j.adaj.2023.02.008

Sheridan, A., O'Sullivan, J., Fisher, J., Dunne, K., & Beck, W. (2020). Responding to institutional climate change in higher education: The evolution of a writing group to group peer mentoring. *Gender and Education, 32*(7), 891–907. https://doi.org/10.1080/09540253.2019.1594705

Shi, L. (2011). Common knowledge, learning, and citation practices in university writing. *Research in the Teaching of English, 45*(3), 308–334. https://www.proquest.com/scholarly-journals/common-knowledge-learning-citation-practices/docview/849266698/se-2

Sutton, R. I., & Staw, B. M. (1995). What theory is not. *Administrative Science Quarterly, 40*(3), 371–384. https://doi.org/10.2307/2393788

Steffen, B. (2025). How hot is the water? In B. Steffen, E. A. Lee, & B. Steffen (Eds.), *Let's talk AI: Interdisciplinarity is a must* (pp. 304–315). Springer. https://doi.org/10.1007/978-3-032-09008-9_32

Taylor & Francis. (2025). AI policy. https://taylorandfrancis.com/our-policies/ai-policy/

Trafford, V., & Leshem, S. (2009). Doctorateness as a threshold concept. *Innovations in Education and Teaching International, 46*, 305–316. https://doi.org/10.1080/14703290903069027

UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000386693

Vidolov, S. (2022). Uncovering the affective affordances of videoconference technologies. *Information Technology & People*, *35*(6), 1782–1803. https://doi.org/10.1108/ITP-04-2021- 0329

Watermeyer, R., Phipps, L., Lanclos, D., & Knight, C. (2024). Generative AI and the automating of academia. *Postdigital Science & Education*, *6*, 446–466. https://doi.org/10.1007/s42438-023-00440-6

Wellington, J. (2013). Searching for 'doctorateness'. *Studies in Higher Education*, *38*(10), 1490–1503. https://doi.org/10.1080/03075079.2011.634901

Wright, A. (2024). Postgraduate supervision in a ChatGPT world: What's next? In *10th International Conference on Higher Education Advances (HEAd'24)*. Valencia, 18–21 June 2024. https://doi.org/10.4995/HEAd24.2024.17244

Yunis, H. (Ed.). (2011). *Plato: Phaedrus*. Cambridge: Cambridge University Press.

CHAPTER THREE

# Redefining Postgraduate Supervision in the Age of AI: Balancing Technology and Human Mentorship


**Edmore Chinhamo[1]**
**Godwin P. Dzvapatsva[2]**
**Courage Matobobo[3]**

AFFILIATIONS

[1]Executive Business Centre, Greenwich Business School, University of Greenwich, London, United Kingdom.
[2]Department of Computing, University of Suffolk, United Kingdom.
[3]Business and Applications Development, Walter Sisulu University, Buffalo City, South Africa.

CORRESPONDENCE

Email: g.dzvapatsva@uos.ac.uk



REFERENCE

Chinhamo, E., Dzvapatsva, G. P., & Matobobo, C. (2026). Redefining postgraduate supervision in the age of AI: Balancing technology and human mentorship. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 38–54). ERRCD Forum. https://doi.org/10.38140/obp4-2026-03





**Abstract:** The advent of technology, particularly the rapid advancement of Artificial Intelligence (AI), is posing significant challenges to traditional models of postgraduate student supervision, ranging from affective mentorship relationships to automated interactions. AI-powered tools such as ChatGPT, Grammarly, DeepSeek, and automated data analysis software provide unprecedented data support to students, thereby enhancing and automating routine tasks. Consequently, the role of supervisors in upholding the fundamental principles of mentoring, such as fostering critical thinking, creativity, and ethical inquiry, is being scrutinised in light of this technological shift. This chapter examines the challenges associated with the incorporation of AI into postgraduate supervision, investigating its impact on intellectual independence, academic integrity, and mentor-mentee dynamics. Through a comprehensive Systematic Literature Review, this conceptual paper identifies strategies for balancing AI-driven efficiencies with human-centred mentoring practices. Additionally, we address ethical considerations, power dynamics, and equity issues that arise within AI-mediated supervision. Our contributions suggest that while AI offers transformative potential, it is essential to preserve the human elements of supervision, empathy, intuition, and the capacity to inspire original thought. This chapter contributes to the ongoing conversation on redefining postgraduate supervision in the digital age, providing actionable insights for supervisors navigating the challenges and opportunities presented by AI.



***Keywords:*** Academic integrity, Artificial Intelligence, critical thinking, ethical inquiry, intellectual independence, mentorship, postgraduate supervision.


## 1. Introduction

In the contemporary academic landscape, universities worldwide are considering policies regarding the ethical and responsible use of Artificial Intelligence (AI), particularly in balancing its use with human effort during learning (Omodan, 2025). Generally defined, artificial intelligence refers to a machine's capacity to reason, communicate, and function autonomously in both known and unfamiliar situations, similarly to a human (Du-Harpur et al., 2020). This means that tasks performed by a human being can also be executed by AI. In the context of this chapter, all the supervisory tasks that a supervisor undertakes during the student supervision process can likewise be performed by AI. These tasks include student supervision and mentorship. The adoption and integration of AI in student supervision, particularly at the postgraduate level, have ushered in new paradigms and reshaped traditional mentorship models. The advent of AI has led to greater confusion about acceptable conduct or codes of behaviour between mentors and mentees during the research process. Consequently, institutions of higher education have mandated that new supervisors undergo training before supervising students, while even qualified supervisors are required to participate in continuing professional development (Mpofu & Madlela, 2024). The current initiative by most higher education institutions to provide courses for student mentors is critical, as supervision involves not merely checking the student's writing but also guiding them through intricate research projects that foster critical thinking and encourage academic growth (Eley & Jennings, 2005), which can be challenging if a mentor lacks exposure.

Before the introduction of AI, it was the mentor's responsibility to shape the academic and professional trajectories of their mentees, providing tailored feedback, emotional support, and career advice (Whitcomb, 2025). In a study by Bouzar et al. (2025) examining how postgraduate students perceive ChatGPT-generated feedback compared to traditional supervisory feedback, the authors found that supervisory input is appreciated for its contextual relevance and level of engagement, while AI-generated feedback is gaining popularity due to its ability to provide quick and consistent evaluations. The literature reviewed indicates that supervisors and mentors may not be readily accessible to their students, and their feedback may not always be clearly noticeable. Paradoxically, engaging with supervisors is a crucial element of learning (Li et al., 2025). In contrast to supervisory feedback, AI-generated feedback is regarded as understandable and readily available whenever needed (Nazaretsky et al., 2024). A balancing act is urgently needed to bridge the human touch with technology. Over the years, postgraduate student supervision has consistently been characterised by a deeply personal and rigorous practice (Ringo, 2025). In this chapter, we examine the intricate nuances of integrating AI into postgraduate supervision while preserving the essential human elements of supervision and mentorship.

Previous studies that have explored the use of AI in student supervision have shown a convergence of views on its usefulness as a tool for mentorship in education. The studies seem to agree with Kulhavy and Wager's (1993) triadic feedback model, as cited by Bouzar et al. (2025), which posits that feedback serves a multilateral function: to keep mentees motivated through response reinforcement, to inform mentees by providing corrective avenues, and to enhance feedback by associating correct responses with prior stimuli. Kulhavy and Wager (1993) suggest that supervision at the postgraduate level extends beyond mere academic guidance to involve the development of mutual trust, understanding the mentee's goals, and providing holistic support. This holistic support entails in-depth intellectual engagement, where ideas are exchanged, challenged, and refined. As previously indicated, this engagement is crucial for developing critical thinking and creativity (Bouzar et al., 2025; Matobobo et al., 2025). Therefore, we posit that the foundation of mentorship is rooted in human connection. Supervisors possess a wealth of experience, wisdom, and empathy that AI cannot replicate. A study by Thong et al. (2025) aimed to investigate how GenAI facilitates the doctoral supervision process, concluding that AI is an effective collaborative tool between the learner and the supervisor. Collaboration in supervision can take various forms, ranging from automated feedback and task management to research mentorship, literature discovery, and collaborative writing and editing (Maor et al., 2016; Lee, 2019). Mentees will spend less time addressing grammatical, structural, and other aesthetic issues in their work. As suggested by Choudhary et al. (2024), when feedback is automated, mentees are freed from the onerous, tedious, and repetitive tasks that are synonymous with traditional mentorship in student research.

Another study by Dai, Lai, Lim, and Liu (2023) examined the impact of ChatGPT, one of the most popular AI tools, on functional enculturation, critical thinking, emancipation, and relationship development during the mentorship process. Their findings reveal that the relationships and roles of mentees and mentors are being transformed by AI, with the mentee transitioning from a mere research apprentice to an independent researcher, while the mentor is tasked with providing strategic direction for the collaboration and research. In the traditional model of student supervision, the mentor typically dominates the relationship, with the mentee following the mentor's instructions. Mentors who have been supervising mentees for a long time may find it difficult to accept the autonomy of their mentees. This is a natural instinct; any change that threatens the sovereignty of mentor power is likely to be resisted. While this is true, some mentors will find the use of AI to be innovative and effective (Choudhary et al., 2024).

AI is proving to be a disruptive tool in student supervision, bringing efficiencies and effectiveness (Halagatti et al., 2023). AI feedback is popular for its speed, consistency, and data-driven insights. It can quickly identify errors in the mentee's work, suggest workable improvements, and provide detailed analyses based on predefined criteria (Thong, Atallah, Islam, Lim, & Cherukuri, 2025). This can be particularly useful for the technical aspects of research, such as statistical analysis or coding. It is evident that the benefits of AI are two-pronged; firstly, it can help the mentee attain independence by utilising AI tools in both the research process and self-management. Secondly, it can assist the mentor in organising the mentoring process effectively and managing their diary properly. However, the impact of physical interaction between a mentee and mentor cannot be underestimated. Face-to-face meetings are essential for detecting non-verbal communication, building mentor-mentee relationships, and facilitating spontaneous, unplanned discussions. If the mentor can detect non-verbal cues during physical meetings, it is easier to manage the mentee's emotions and offer reassurance where appropriate. Finding a balance that integrates AI mentorship with traditional supervision is critical to addressing paradigm shifts in student mentorship.

In other arguments, Bouzar et al. (2025) suggest that the traditional mentee-mentor relationship, which is neither influenced nor enhanced by AI, is effective in promoting engagement. The mentee is able to interact with the mentor in real time. The premise is that apprenticeships are the best way to learn; in this context, the mentee acts as an apprentice to the mentor, who oversees the entire research process. In the traditional model of the mentee-mentor relationship, it often took a long time for the mentor to provide feedback to the mentee. Planning was also difficult, as the mentee had to physically travel to meet the mentor on campus. Mentors hold more power, and as a result, the relationship between mentor and mentee can sometimes be toxic (Dai et al., 2023; Paulsen & Schmidt-Crawford, 2017). Aymen and Zakarya (2024) argue that AI-generated solutions lack originality. AI tools have gained fame for their speed and accuracy in problem-solving. However, concerns about relying too heavily on AI are likely to impact the development of cognitive skills required in the current job market, including critical thinking and problem-solving (Chang et al., 2024; Wright, 2024). The growing use and overdependence on AI may produce cohorts of graduates who could face challenges in utilising their mental faculties to solve problems or engage in meaningful academic debate.

## 1.1 Problem statement

The advent of AI is rapidly disrupting postgraduate supervision in the higher education sector. As a result, many institutions are facilitating its adoption and use to prepare for the fast-approaching market changes (Hutson et al., 2022). While some universities are accelerating the use of AI, others are conflicted by the complexities of its responsible and ethical application (Slimi & Carballido, 2023). Although the adoption of AI may seem plausible at face value, the reality tells a different story. Traditional models, which emphasised close, personalised mentorship between mentees and mentors, are being challenged by the integration of AI tools that can outline research designs, author articles, analyse data, and manage projects. While AI offers efficiency and easy access to knowledge, its adoption poses a threat to the human factors essential for critical thinking and ethical reasoning. The current discourse on the use of AI in higher education lacks substantive frameworks for balancing its use with human mentorship. There is a risk of eroding the gains, depth, and quality of postgraduate research that has been achieved over the years through traditional human mentorship. Furthermore, these conflicting paradigms present challenges to policymakers and higher education institutions regarding the direction to take on AI adoption and its use.

### *1.1.1 Research question*

The study is guided by the following research question:
*How can postgraduate supervision models be restructured to balance AI tools with human mentorship?*

### *1.1.2 Theoretical underpinnings: Connectivism approach*

This study is grounded in Connectivist Learning Theory, which was introduced by George Siemens in 2004 and published in 2005. Connectivism offers a theoretical framework for understanding learning in technology-mediated environments, characterised by interconnected nodes, dynamic networks, complexity, and openness (Siemens, 2005; Goldie, 2016). The authors argue that learning occurs when learners connect, share, interact, and collaborate with members of a network community. AI has the capacity to stimulate self-learning and collaboration with others, which determines and redefines the extent of the supervisor's involvement in the student's work. Siemens (2004) posits that, under connectivism, the knowledge repository continues to grow due to the extensive use of technology, which serves as a source of up-to-date information. The arguments presented by both Siemens (2005) and Goldie (2016) are predicated on the importance of networks and distributed knowledge. Knowledge is no longer confined to the minds of individuals but is distributed across technological networks and databases. However, the authors also contend that connectivism may not be an appropriate standalone approach, as it focuses on how learners utilise digital technologies and interact with peers to acquire and build knowledge. In the context of the current study, AI can be viewed as a node in the student's learning network, while emphasising the supervisor's role in curating and validating learning paths. The implementation of AI in supervision can be seen as chaotic.

Connectivism has been criticised by Verhagen (2006) for its philosophical deficiencies. Although the advent of technology is shaping how humans can acquire knowledge through databases or connections, proponents of connectivism do not adequately explain how people 'come to know' (Clara & Barbera, 2013). Kop and Hill (2008) present a similar argument, suggesting that knowledge-making and transfer have not been sufficiently explained within the connectivist perspective. The availability of knowledge repositories, which are largely accessible across various networks, may diminish the importance of knowledge-making as long as learners know where to find the information when needed. When learners are aware of when to seek knowledge, interaction with tutors or supervisors is limited if they can navigate through knowledge repositories or learning networks to find what they are looking for. The assumption that where there is a network, there is learning is unsustainable; there are instances of learning where the supervisor's role is indispensable. The empathetic nature of human supervisors in addressing student queries makes AI a more complementary tool than a definitive one for learning.

## 2. Methodology

This study employed a qualitative approach and a systematic literature review (SLR) to explore how AI is influencing mentor-mentee relationships and supervision practices in postgraduate education. A qualitative approach was chosen because it is best suited to capturing the depth and complexity required to redefine supervision models in an AI-infused academic environment. The systematic literature review rigorously analyses existing research on AI's role in postgraduate supervision, ensuring a comprehensive and unbiased synthesis of diverse perspectives. The SLR methodology enables the identification of key trends, gaps, and best practices, providing a strong evidence base for redefining mentorship in the AI era. According to Okoli and Schabram (2010), the evaluation of scholarly works ensures transparency, reproducibility, and actionable insights for academia and policymakers.

### 2.1 Search strategy

A systematic search was conducted across three electronic databases: Scopus, Web of Science, and Google Scholar. These databases were selected for their comprehensive coverage of peer-reviewed literature in education, technology, and interdisciplinary research relevant to postgraduate supervision. To ensure consistency across the databases, the search strategy combined key concepts related to postgraduate education, supervision, mentorship, and artificial intelligence. Boolean operators were used to refine and broaden the search where appropriate.

For Scopus, ERIC, and Web of Science, the following search string was applied: (postgraduate) AND (supervision OR mentorship) AND (AI OR "artificial intelligence"). This search yielded 15 articles in Scopus and 48 articles in Web of Science, all published between 2020 and 2025. For Google Scholar, a more specific phrase-search approach was employed to manage the higher volume of indexed material: ("postgraduate supervision") AND ("artificial intelligence"). This search produced 415 results. All retrieved articles were exported for screening, and duplicates were removed prior to the eligibility assessment stage.

#### *2.1.1 Inclusion criteria*

The study was guided by the following inclusion criteria:

- Articles focusing on postgraduate supervision or mentorship
- Articles focusing on the use of AI technologies in postgraduate supervision
- Articles focusing on the integration, benefits, challenges, or risks of AI technologies in postgraduate supervision
- Articles are peer-reviewed
- Articles are written in the English language.
- Full articles are available online

The study generated a total of 478 articles across the three databases, as shown in Figure 1. After removing 32 duplicate articles, 446 articles remained. An initial screening of these 446 articles was conducted using their abstracts and titles. From this screening, 429 articles were excluded for not meeting the inclusion criteria, leaving only 17 articles that met the criteria. A detailed analysis of these 17 articles was performed. Five articles were excluded because they did not focus on AI technologies, and two were excluded as they did not pertain to postgraduate supervision. Consequently, only 10 articles met the inclusion criteria. As the studies were insufficient for the research, additional studies were sought from other Google search engines using the search string (postgraduate) AND (supervision OR mentorship) AND (AI OR "artificial intelligence”). Five studies met the inclusion criteria and were added to the research. Ultimately, 15 studies were included in the final analysis.

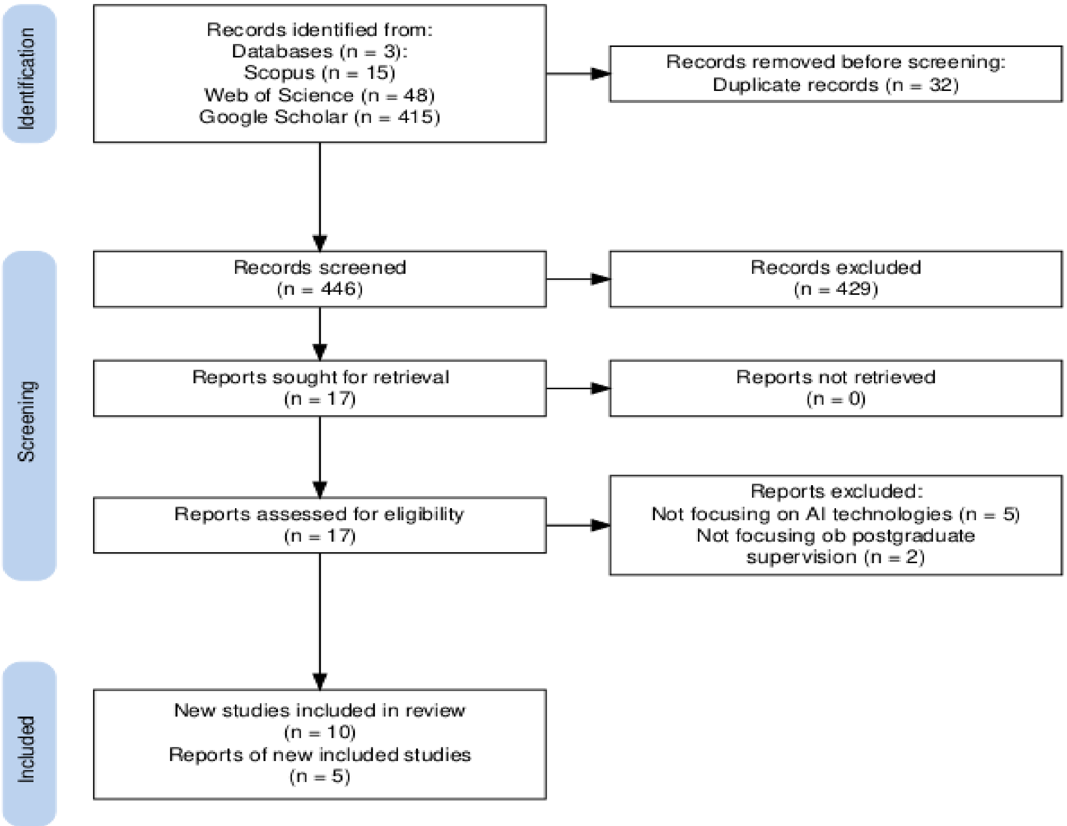



***Figure 1:*** *PRISMA flow diagram*

The 15 articles included in the study were analysed using ATLAS.ti. Thematic analysis was conducted following the six steps recommended by Braun & Clarke (2006). The first step involved the authors familiarising themselves with the data. The second step entailed developing initial codes relevant to the study. The third step consisted of identifying themes by merging the generated codes. The fourth step focused on refining the themes. The fifth step required describing and naming the themes by generating distinct meanings for each one. The sixth step involved reporting on the gathered themes, which are presented in the subsequent section.

## 3. Presentation of Results

The results are presented in the form of themes. The chapter generated five themes and 23 codes, as shown in Table 1.

***Table 1:*** *Themes and codes*

| Theme | Codes |
|---|---|
| Power dynamics and trust in AI-mediated Supervision | AI as disruptor<br>Fear of academic dishonesty<br>Shifting power balances<br>Need for transparency<br>Ethical vigilance |
| Role of transformation in the AI era | AI handles routine tasks |

| | |
|---|---|
| | Supervisors focus on strategy<br>Students gain autonomy<br>AI as "third wheel" vs collaborator<br>Virtual mentoring |
| Implementation challenges and systemic barriers | Lack of institutional policies<br>Need for training<br>Ethical concerns<br>Cultural adaptation needs |
| Optimal integration models | AI for efficiency<br>Humans for quality<br>Task-technology fit<br>Critical engagement |
| Complementing Human Support and AI in Mentoring | Emotional support<br>Contextual understanding<br>Critical thinking<br>Bias mitigation<br>Ethical oversight |

### 3.1 Power dynamics and trust in AI-mediated supervision

This theme captures the fundamental tensions that arise as AI becomes embedded in postgraduate supervision, reshaping long-standing power relations, trust expectations, and supervisory roles. The integration of AI tools, particularly conversational systems such as ChatGPT, has disrupted traditional postgraduate supervision by altering the power dynamics and trust between students and supervisors. Studies show that students fear being penalised for using AI tools, while supervisors are concerned about reputational risks, compromised academic integrity, and the possibility of students submitting work without attribution (Harding & Boyd, 2024; Wright, 2024). As a result, AI frequently becomes a concealed “third wheel” in the supervision process rather than a transparently negotiated support tool (Wright, 2024).

Evidence across the reviewed studies demonstrates that covert AI use significantly alters power dynamics. Boyd and Harding (2025) reveal that doctoral students increasingly use AI as an “invisible tutor” or “safety net” to avoid criticism, effectively shifting epistemic authority away from supervisors and creating, in their view, a “sinister power behind the scenes.” This hidden reliance complicates trust and undermines the relational foundation of supervision, particularly because supervisors are often unaware of the extent to which AI is used. Mbodila (2025) similarly notes a shift in the supervisor’s role from primary knowledge provider to facilitator of critical AI use, which diffuses traditional authority and raises new integrity concerns.

Transparency emerges as a central mitigating principle. Iatrellis et al. (2025) argue that AI involvement in supervision must be “visible and accountable,” warning that non-disclosure threatens ethical authorship practices and relational trust. This aligns with broader ethical concerns highlighted by Köbis and Mehner (2021), who emphasise that AI systems introduce opacity and algorithmic influence into mentoring relationships, contexts that require exceptionally high levels of trust. Jensen et al. (2025) add further nuance by showing how AI redistributes agency in feedback processes. Students gain more control during chatbot interactions, as chatbots follow the user's agenda, whereas supervisors exercise traditional authority by redirecting attention and shaping scholarly engagement.

Overall, these dynamics highlight an emerging landscape of ethical vigilance (Dai et al., 2023), where supervisors must monitor issues such as fabricated references, biased outputs, or inappropriate reliance on AI (Segooa et al., 2025), all while operating without clear institutional policies or shared norms. Collectively, the studies indicate that the introduction of AI into postgraduate supervision reconfigures trust, destabilises established hierarchies, and requires deliberate strategies of transparency, guidance, and ethical oversight to maintain a healthy supervisory relationship.

### 3.2 Role of transformation in the AI era

The results indicate that AI integration is reshaping postgraduate supervision by altering the roles of both students and supervisors. AI tools, such as ChatGPT, enhance student autonomy by supporting self-directed learning and assisting with routine tasks, including grammar checks and idea generation (Sim et al., 2023; Dai et al., 2023). This shift allows supervisors to concentrate on higher-order academic guidance, such as refining research questions, ensuring methodological rigour, and fostering critical engagement, rather than spending time on technical editing (Dai et al., 2023; Mbodila, 2025). The transformation of supervisory roles is further emphasised by Mbodila (2025), who argues that supervisors now function as "AI literacy mentors," responsible for promoting ethical AI use, maintaining academic integrity, and preserving the human element in supervision. The supervisory relationship increasingly resembles a balanced triad, where the student, supervisor, and AI each contribute distinct expertise (Dai et al., 2023; Iatrellis et al., 2025). Iatrellis et al. (2025) formalise this shift through the tripartite mentoring model, in which AI provides scalable support, supervisors offer domain expertise and ethical oversight, and students engage critically with AI outputs. From the student perspective, AI is sometimes perceived as a surrogate mentor, offering emotional, exploratory, and pedagogic support, thereby reshaping expectations of the supervisor (Boyd & Harding, 2025). Jensen et al. (2025) further distinguish the complementary strengths of humans and AI, noting that chatbots excel at immediate, task-focused assistance, while supervisors remain essential for developmental, relational, and context-rich feedback. Collectively, these insights demonstrate that, although AI supports emerging hybrid and virtual supervision practices, it cannot replace the supervisory functions founded on pedagogical intention, identity development, and care (Dai et al., 2023; Jensen et al., 2025).

### 3.3 Implementation challenges and systemic barriers

AI in postgraduate supervision faces multifaceted implementation challenges and systemic barriers that span policy, ethics, training, and institutional culture. Studies reveal three institutional shortcomings. First, persistent policy gaps leave supervisors without clear standards for evaluating AI-assisted work (Wright, 2024). Second, inadequate training limits supervisors' readiness to integrate AI meaningfully into their pedagogy (Cowling et al., 2023). Third, cultural adaptation challenges, especially in Global South contexts where students already face resource constraints, further magnify inequitable access to AI tools (Segooa et al., 2025). These barriers intersect with detection difficulties, as institutions often lack the tools to detect sophisticated AI use, resulting in enforcement dilemmas and highlighting a sector struggling with AI integration (Wright, 2024). Further evidence shows that the covert and unacknowledged use of AI, driven by unclear institutional guidelines and the stigma surrounding AI, creates fear, confusion, and a reluctance among academics to openly discuss best practices (Boyd & Harding, 2025). Ethical and practical risks, such as over-reliance on AI, academic integrity violations, limited AI literacy, and the absence of robust supervisory policies, all contribute to a substantial implementation gap between identifying problems and developing institutional responses (Mbodila, 2025). Ethical concerns extend to data privacy, algorithmic bias, and the wide variability in prompt-engineering competence, which restricts the effectiveness and reliability of AI-supported feedback (Iatrellis et al., 2025; Jensen et al., 2025). Broader systemic tensions further complicate adoption, including conflicting institutional priorities, financial constraints, and disciplinary divides, where educators may overlook technical risks while AI developers may underestimate pedagogical ethics (Köbis & Mehner, 2021). Collectively, these findings demonstrate that while AI holds potential value for postgraduate supervision, its integration is hindered by intertwined policy, ethical, technical, and cultural barriers that institutions have yet to adequately address.

### 3.4 Optimising hybrid supervision models

To utilise AI effectively, it is essential to tailor the technology to the specific task at hand. For example, the use of Elicit can streamline literature discovery while allowing students to focus on deeper conceptual development (Segooa et al., 2025). However, AI outputs must be approached critically, as studies show that students should not accept responses at face value due to the risk of inaccuracies (Sim et al., 2023) and fabricated citations, which require human oversight and correction (Bouzar et al., 2023). Ethical safeguards remain central, particularly in preventing bias and ensuring responsible use (Cowling et al., 2023). Existing evidence also suggests that AI can help balance the supervisory workload by providing timely support for routine tasks (Iwashokun & Ade-Ibijola, 2022; Serek & Zhaparov, 2024). However, scholars caution against overreliance, which may make the research process feel overly mechanistic and detached (Dai et al., 2023).

Beyond tool-use guidance, several studies propose structured frameworks for integrating AI into supervision. Iatrellis et al. (2025) introduce the Tripartite Mentoring Model, grounded in principles such as complementary roles, collaborative dialogue, transparency, AI literacy, and ongoing evaluation, framing supervision as a coordinated interaction between student, supervisor, and AI. Similarly, Boyd and Harding (2025) argue for shifting from a covert "third wheel" model to openly giving AI a "seat at the table," emphasising the need to avoid hidden, unacknowledged AI use. Mbodila (2025) recommends a hybrid model in which AI functions as a supplementary research assistant while human supervisors retain responsibility for nurturing creativity, critical thinking, and methodological rigour. From an ethical design perspective, Köbis and Mehner (2021) propose an interdisciplinary checklist that integrates AI ethics (e.g., transparency, robustness) with mentoring ethics (e.g., confidentiality, beneficence), encouraging the adoption of value-sensitive and stakeholder-inclusive AI. At the practical user level, Jensen et al. (2025) provide guidance such as using chatbots to prepare for meetings, ask questions, validate understanding, and translate complex material, while stressing that relational pedagogies must guide integration and that AI systems cannot assume supervisory responsibility.

### 3.5 Complementing human support and AI in mentoring

This theme affirms the irreplaceable role of human expertise in postgraduate supervision while highlighting the complementary value of AI. Studies consistently show that human supervisors provide emotional support, psychosocial guidance, relational depth, and contextualised scholarly judgment—qualities that AI cannot replicate (Bouzar et al., 2025; Cowling et al., 2023). Human supervisors also engage in critical thinking to mitigate algorithmic bias and ensure ethical decision-making, positioning them as guardians of academic integrity in an AI-mediated research environment (Harding & Boyd, 2024; Cowling et al., 2023). Several studies explicitly emphasise that AI should augment, not replace, human mentorship. Iatrellis et al. (2025) describe a tripartite model in which AI handles scalable, time-consuming tasks, such as generating initial recommendations, thereby freeing supervisors to focus on high-level intellectual guidance and ethical oversight, while students synthesise insights from both sources. Similarly, Mbodila (2025) stresses that AI remains a supplementary tool; the emotional support, trust-building, creativity nurturing, and personalised academic feedback provided by human supervisors are irreplaceable. Empirical evidence from Boyd and Harding (2025) suggests that although AI can temporarily fill certain support gaps, such as providing immediate clarification or reassurance, it does so only superficially. Instead, institutions should use this insight to reinforce areas where human engagement is most needed. Jensen et al. (2025) further reinforce this complementary relationship, demonstrating that chatbots are effective for instrumental, task-focused assistance, whereas supervisors provide developmental, identity-shaping, and relational feedback. Collectively, these studies underline that AI can enhance supervision efficiency, but the essence of mentorship—the human connection—remains central and indispensable.

## 4. Discussion of Findings

This study aimed to explore how postgraduate supervision models can be restructured to balance the use of AI tools with human mentorship. This scoping review identified five consistent themes across the literature:

- Power dynamics and trust in AI-mediated supervision
- Role of transformation in the AI era
- Implementation challenges and systemic barriers
- Optimising hybrid supervision models

- Complementarity of human support and AI in mentoring

These themes reflect a complex and evolving landscape where AI is reshaping traditional supervisory relationships, roles, and systems, prompting both opportunities and ethical considerations.

AI in supervisory roles raises concerns about power and trust. The reviewed literature has shown that trust in such systems largely depends on transparency, fairness, and ethical design (Kulhavy & Wager, 1993). Human oversight is thus essential to ensure that AI enhances rather than undermines the relationship between students and mentors at postgraduate levels (Jensen et al., 2025). Undoubtedly, students are more likely to trust AI, while supervisors, who may already be familiar with concerns about the algorithmic bias of models as reported in academic circles, may attempt to trust AI within limits. Without careful safeguards, the supervisory relationship risks becoming impersonal or coercive.

Our reviewed articles have demonstrated the potential of AI to transform the role, with most routine processes now achievable through the technology (Dai et al., 2023; Mbodila, 2025). Although this transformation has been noted, it has also created new demands for supervisors to acquire digital skills to critique AI output. This study anticipates potential tensions between management and supervisors, particularly regarding workload expectations, as management may assume that AI will enable significantly higher output. There is a growing sense of uncertainty surrounding AI integration, as institutions are hastily developing policies in response—an approach that risks premature implementation and insufficient stakeholder engagement (Wright, 2024; Matobobo et al., 2025). Notably, divergent perceptions of AI integration are emerging across academic platforms, including conferences and peer-reviewed journals, reflecting a lack of consensus within the scholarly community (Boyd & Harding, 2025).

While our findings suggest that AI is for efficiency and humans for quality, resulting in a balance, the quality teams in most universities may have different perceptions, adding another layer of complexity to the hybridisation process. The findings underscore the importance of striking a balance between human and AI support for emotional well-being, which is particularly crucial in supervisory contexts (Iatrellis et al., 2025). There are times when the mentee requires emotional support during the supervision process that AI cannot offer (Mbodila, 2025). An aspect that did not emerge in the findings is the position of contract writers. This study suggests that AI has the potential to alleviate the growing, yet often unspoken, challenge of contract writing in academic institutions, enabling students to complete tasks that were previously difficult to accomplish within limited timeframes (Sim et al., 2023; Dai et al., 2023). Given the rapid advancement of AI technology, current limitations may prove to be short-lived, as ongoing developments are likely to address many of the tool's perceived shortcomings in the near future. In summary, the lack of compassion in AI limits its ability to comprehend human emotions, often resulting in harmful or insensitive outcomes. This shortfall undermines trust and poses a barrier to the technology's broader acceptance and effectiveness.

## 5. Conclusions

This chapter examined the potential restructuring of postgraduate supervision models to achieve a balance between the integration of artificial intelligence (AI) tools and the enduring value of human mentorship. Guided by connectivist theory, the study conceptualised AI as an additional "node" within the supervisory relationship, rather than as a replacement for the supervisor. The findings from the systematic review reveal that postgraduate supervision is currently undergoing a transitional phase, characterised by shifting power dynamics, evolving roles of supervisors and students, and an increasing emphasis on trust and ethical responsibility within AI-mediated academic environments.

Across the reviewed literature, five themes emerged: power dynamics and trust in AI-mediated supervision, the role of transformation in the AI era, implementation challenges and systemic barriers, the optimisation of hybrid supervision models, and the complementarity of human support and AI in mentoring. Collectively, these themes underscore that while AI offers efficiency, scalability, and on-demand academic support, it cannot replicate the relational, developmental, and ethical dimensions of supervision. Human supervisors remain central to providing psychosocial support, exercising scholarly judgment, and safeguarding academic integrity. The study, therefore, contributes to both practice and theory by demonstrating that effective postgraduate supervision in the AI era relies on hybrid models that intentionally integrate AI capabilities with human-centred supervisory practices.

### 5.1 Recommendations

Based on the findings of this study, the following recommendations are proposed for higher education institutions, supervisors, and policymakers:

- Institutions could formally adopt hybrid supervision models that recognise AI as a complementary support tool while preserving the primacy of human mentorship. Such models should clearly define the respective roles of supervisors, students, and AI systems to prevent over-reliance on technology and to maintain accountability.
- Supervisors could be encouraged to participate in Continuous Professional Development (CPD) programmes focused on the pedagogically sound and ethical integration of AI in supervision. These programmes should address issues such as bias detection, ethical oversight, feedback practices, and managing power dynamics in AI-enhanced supervisory relationships.
- Higher education institutions must develop and enforce clear policies and guidelines governing the responsible use of AI in postgraduate research and supervision. These policies should clearly articulate the uses of AI, outline expectations for transparency, and establish mechanisms for maintaining academic integrity.
- Future research should expand beyond English-language studies to reduce geographic and linguistic bias and to capture perspectives from underrepresented regions. Empirical

studies examining the long-term effects of AI-assisted supervision on student agency, trust, and academic identity would further strengthen the evidence base for informed policy and practice.

## 6. Declarations

**Funding:** This research did not receive any external funding.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Use of Artificial Intelligence:** The current work was created with the assistance of artificial intelligence technologies (ChatGPT-4 and Grammarly) to assist with refining language for clarity, as confirmed by the authors.

## References

Aymen, D., & Zakarya, B. (2024). *The influence of artificial intelligence on students' critical thinking: A case study of third year students of English at the University Center of Mila.* Abdelhafid Boussouf University-Mila.

Bouzar, A., El Idrissi, K., Ghourdou, T., & Ali, N. (2025). Supervisory feedback vs. AI: A comparative study on postgraduate student satisfaction. *ISRG Journal of Arts, Humanities & Social Sciences, 3*(1), 465–474. https://doi.org/10.5281/zenodo.14862530

Boyd, P., & Harding, D. (2025). Generative AI: Reconfiguring supervision and doctoral research. *Buildings & Cities, 6*(1), 1-23.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Chang, C. N., Hui, J., Justus-Smith, C., & Wang, T. W. (2024). Navigating STEM careers with AI mentors: A new IDP journey. *Frontiers in Artificial Intelligence, 7,* 1461137. https://doi.org/10.3389/frai.2024.1461137

Choudhary, P., Sharma, K., Sharma, K., Borasi, M., Bhargava, P., Ali, I., & Rehman, K. (2024). Enhancing mentorship through technology: A comprehensive review of current practices and future directions. *International Journal of Multidisciplinary Research and Growth Evaluation, 5*(6), 634–645.

Clara, M., & Barbera, E. (2013). Three problems with the connectivist conception of learning. *Journal of Computer Assisted Learning, 30*(3), 197–206.

Cowling, M., Crawford, J., Allen, K. A., & Wehmeyer, M. (2023). Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research supervision. *Australasian Journal of Educational Technology, 39*(4), 89–103.

Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology, 39*(4), 74–88.

Du-Harpur, X., Watt, F. M., Luscombe, N. M., & Lynch, M. D. (2020). What is AI? Applications of artificial intelligence to dermatology. *British Journal of Dermatology, 183*(3), 423–430. https://doi.org/10.1111/bjd.18880

Eley, A., & Jennings, R. (2005). *Effective postgraduate supervision: Improving the student/supervisor relationship.* McGraw-Hill Education.

Goldie, J. G. S. (2016). Connectivism: A knowledge learning theory for the digital age? *Medical Teacher, 38*(10), 1064–1069.

Jensen, L.X., Bearman, M., Boud, D., & Konradsen, F. (2025). Feedback encounters in doctoral supervision: The role of generative AI chatbots. *Assessment & Evaluation in Higher Education,* 1–14.

Halagatti, M., Gadag, S., Mahantshetti, S., Hiremath, C. V., Tharkude, D., & Banakar, V. (2023). Artificial intelligence: The new tool of disruption in educational performance assessment. *Smart analytics, artificial intelligence and sustainable performance management in a global digitalised economy*, 261–287.

Harding, D., & Boyd, P. (2024). Generative AI and PhD supervision: A covert third wheel or a seat at the table? *ARCOM Conference 2024: Looking back to move forward 40th conference.* https://ucem.repository.guildhe.ac.uk/id/eprint/125/1/GenAI_registry.pdf

Hutson, J., Jeevanjee, T., Graaf, V. Vander, Lively, J., Weber, J., Weir, G., Arnone, K., Carnes, G., Vosevich, K., Plate, D., Leary, M., & Edele, S. (2022). Artificial intelligence and the disruption of higher education: Strategies for integration across disciplines. *Creative education*, *13*(12), 3953–3980. https://doi.org/10.4236/ce.2022.1312253

Iatrellis, O., Bania, A., Samaras, N., Kosmopoulou, I., & Panagiotakopoulos, T. (2025). ChatGPT in doctoral supervision: Proposing a tripartite mentoring model for AI-assisted academic guidance. *International Journal of Doctoral Studies, 20(9),* 1-26.

Iwashokun, O., & Ade-Ibijola, A. (2022). Structural vetting of research proposals: Problematisation and solving with artificial intelligence. *Proceedings of International Conference on Information Systems and Emerging Technologies.* https://ssrn.com/abstract=4331475

Kop, R., & Hill, A. (2008). Connectivism: Learning theory of the future or vestige of the past? *International Review of Research in Open and Distributed Learning, 9*(3), 1-13.

Kulhavy, R. W., & Wager, W. (1993). Feedback in programmed instruction: Historical context and implications for practice. *Interactive Instruction and Feedback*, 320.

Köbis, L., & Mehner, C. (2021). Ethical questions raised by AI-supported mentoring in higher education. *Frontiers in Artificial Intelligence, 4*, 624050.

Lee, A. (2019). *Successful research supervision: Advising students doing research.* Routledge.

Li, Y., Xu, W., & Chen, J. (2025). PhD student-supervisor relationship and its impacts: A perspective of the interpersonal relationship model. In *Frontiers in Education* (Vol. 10, p. 1570137). Frontiers Media SA.

Matobobo, C., Ncube, P. D. N., Ngesimani, N. L., Dzvapatsva, G. P., & Chinhamo, E. (2025). Enhancing computational thinking and problem-solving in programming education through generative AI: A scoped review. *2025 IEEE Global Engineering Education Conference* (EDUCON) (pp. 1-9). IEEE.

Maor, D., Ensor, J.D., & Fraser, B.J. (2016). Doctoral supervision in virtual spaces: A review of research of web-based tools to develop collaborative supervision. *Higher Education Research & Development, 35*(1), 172-188.

Mbodila, M. (2025). Postgraduate supervision in the age of ChatGPT: Redefining the role of supervisors. *Foundations and Trends® in Information Systems, 9*(2), 60-68.

Mpofu, P., & Madlela, B. (2024). An inquiry into experiences of supervisors and supervisees in conducting research supervision in higher education institutions in Eswatini. *E-Journal of Humanities, Arts and Social Sciences, 5*(8), 1454-1473.

Nazaretsky, T., Mejia-Domenzain, P., Swamy, V., Frej, J., & Käser, T. (2024, September). AI or human? Evaluating student feedback perceptions in higher education. In *European conference on technology enhanced learning* (pp. 284-298). Cham: Springer Nature Switzerland.

Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *Sprouts: Working Papers on Information Systems*, *10*(26), 1-14.

Omodan, B. I. (2025). Redefining the role of supervisors in the era of artificial intelligence: Implications for hybrid postgraduate research governance. *Cogent Education*, *12*(1), 2536534.

Paulsen, T. H., & Schmidt-Crawford, D. A. (2017). Enhancing student teacher supervision through hybridization: Adding e-supervision to the mix. *Journal of Agricultural Education*, *58*(2), 166–179. https://doi.org/10.5032/jae.2017.02166

Ringo, D. S. (2025). When relationships make a difference! Supervisor support and PhD students' academic performance in the journey: The moderating effect of supervisor-student relationship quality. *Cogent Education*, *12*(1), 2479403.

Segooa, M. A., Modiba, F. S., & Motjolopane, I. (2025). Generative artificial intelligence tools to augment teaching scientific research in postgraduate studies. *South African Journal of Higher Education*, *39*(1), 294-314. https://doi.org/10.20853/39-1-6275

Serek, A., & Zhaparov, M. (2024). Optimizing preference satisfaction with genetic algorithm in matching students to supervisors. *Applied Mathematics and Information Sciences, 18*(1), 133–138. https://doi.org/10.18576/amis/180114

Siemens, G. (2005). *Connectivism: A learning theory for the digital age.* https://www.itdl.org/Journal/Jan_05/article01.htm

Sim, K. N., Northcote, M., & Ping Lim, C. (2023). Technology-enabled undergraduate and postgraduate research supervision. *Australasian Journal of Educational Technology, 39*(4), 1–5.

Slimi, Z., & Carballido, B. V. (2023). Navigating the ethical challenges of artificial intelligence in higher education: An analysis of seven global AI ethics policies. *TEM Journal, 12*(2), 590–602. https://doi.org/10.18421/TEM122-02

Thong, C. L., Atallah, Z., Islam, S., Lim, W., & Cherukuri, A. K. (2025). AI-powered tools for doctoral supervision in higher education: A systematic review. *Journal of Information & Knowledge Management*, 25(3), 1–20. https://doi.org/10.1142/S0219649225300018

Verhagen, P. (2006). Connectivism: A new learning theory. http://elearning.surf.nl/elearning/english/3793

Whitcomb, T. (2025, October 25). The role of mentorship in academic and professional growth of students. *Hispanic Fund*. https://hispanic-fund.org/the-role-of-mentorship-in-academic-and-professional-growth-of-students/

Wright, A. (2024). Postgraduate supervision in a ChatGPT world: What's next? *International Conference on Higher Education Advances*, 480–487. https://doi.org/10.4995/HEAd24.2024.17244

CHAPTER FOUR

# Artificial Intelligence and the PhD: Navigating Doctoralness in the Digital Age

**Winter Sinkala**[1]
**Israel Kariyana**[2]

AFFILIATIONS

[1]Department of Mathematical Sciences and Computing, Faculty of Natural Sciences, Walter Sisulu University, Mthatha, South Africa.
[2]Continuing Professional Teacher Development, Faculty of Education, Walter Sisulu University, Mthatha, South Africa.

CORRESPONDENCE

Email: wsinkala@wsu.ac.za

REFERENCE

Sinkala, W., & Kariyana, I. (2026). Artificial intelligence and the PhD: Navigating doctoralness in the digital age. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 55–73). ERRCD Forum. https://doi.org/10.38140/obp4-2026-04



**Abstract:** The doctorate has long been regarded as the pinnacle of higher educational attainment, demanding originality, critical inquiry, and the capacity to generate new knowledge—qualities collectively referred to as doctoralness. In the early twenty-first century, doctoral education is undergoing transformation due to the increasing prevalence of artificial intelligence (AI) tools in research design, data analysis, academic writing, and supervisory practices. This chapter examines the intersection of AI with the nature and practice of doctoralness. We begin by clarifying the historical and conceptual foundations of doctoralness as an intellectual and identity-forming endeavour that extends beyond mere technical research skills. Subsequently, we explore the evolving landscape of the PhD as candidates, supervisors, and institutions adopt AI-enabled tools for literature synthesis, multilingual writing support, modelling, and personalised feedback. While these tools promise efficiency, inclusivity, and new modes of collaboration, they also pose risks—such as over-reliance, erosion of critical judgment, breaches of academic integrity, and the widening of inequities between well-resourced and under-resourced contexts. Drawing on global literature and examples from South Africa and the Global South, this chapter discusses strategies for safeguarding doctoralness through supervisor professional development, institutional AI literacy frameworks, and policies grounded in ethical and epistemic justice. We argue that the responsible integration of AI can enrich rather than diminish doctoral education when guided by human criticality and robust scholarly norms. The chapter concludes with recommendations for future directions in AI-infused doctoral training within a digitally mediated knowledge society.

***Keywords:*** Artificial intelligence, doctoralness, ethics of AI, human–AI collaboration, postgraduate supervision, research integrity.

## 1. Introduction

Doctoral education has historically signified entry into a community of scholars whose core attributes include independence of thought, originality of contribution, and the capacity to interrogate and advance disciplinary knowledge (Trafford & Leshem, 2009; McKenna & van Schalkwyk, 2024). This distinctive constellation of qualities—often described as doctoralness—embodies a scholarly identity marked by intellectual autonomy, methodological rigour, ethical awareness, and reflexive judgement (Mowbray & Halse, 2010; Wisker, 2012). In recent decades, the meaning of doctoralness has been reframed in response to the massification of doctoral enrolments, demands for societal impact, and the diversification of doctoral career pathways

(Sarrico, 2022; Backhouse, 2009). Guerin et al. (2015) advocate for moving beyond the traditional apprenticeship model towards conceptions of doctoral learning as collaborative, situated practice that equips graduates for complex, digitally mediated knowledge work.

A defining feature of today's knowledge economy is the pervasive influence of artificial intelligence. AI-powered systems—ranging from machine-learning algorithms for data modelling to generative language models for writing support—are rapidly permeating the research workflow (Zawacki-Richter et al., 2019; Holmes et al., 2019). For doctoral researchers, this means that tools capable of summarising vast literatures, translating multilingual texts, suggesting analytical pipelines, or drafting prose are now readily accessible—often at little or no cost.

The growing ubiquity of such tools raises urgent questions about how doctoralness can be cultivated and recognised in an era where aspects of scholarly labour may be partially automated. Does reliance on AI diminish the candidate's demonstration of originality and critical thinking, or can it amplify these qualities by freeing cognitive resources for higher-order reasoning (Nguyen et al., 2024)? How should supervisors balance efficiency gains from AI with their responsibility to foster deep conceptual engagement, methodological understanding, and responsible research conduct (Wisker, 2012)? These questions are particularly salient in the Global South, where uneven access to advanced digital infrastructure risks exacerbating existing inequities in doctoral training (Akala, 2021; Dlamini & Ndzinisa, 2025). At the same time, well-implemented AI tools have the potential to reduce barriers to entry—for example, by supporting students who write in additional languages or by automating routine formatting and citation tasks, thereby enabling more equitable participation in global research dialogues (UNESCO, 2023).

This chapter situates the debate at the intersection of doctoral-level scholarly formation and the accelerating integration of AI into research practice. It traces the historical development of the concept of doctoralness, outlines the evolving technological landscape of the PhD, weighs the promises and perils of AI for doctoral training, and proposes strategies for safeguarding the epistemic integrity of the degree.

### 1.1 The nature of doctoralness

The term "doctoralness" encapsulates the distinctive characteristics of the doctorate: originality, intellectual autonomy, methodological rigour, ethical awareness, and the capacity to generate knowledge that withstands critical scrutiny (Trafford & Leshem, 2009; McKenna & van Schalkwyk, 2024). It reflects both the actions of doctoral candidates and their evolution into scholars who contribute new insights.

Historically, the doctorate has evolved from medieval teaching licences to the Humboldtian research model, which emphasises independent inquiry (Sarrico, 2022). Barnacle and Dall'Alba

(2014) assert that doctoral education is a practice that shapes the being and thinking of the scholar, rather than merely a form of technical training. Traditional apprenticeship-style supervision has been critiqued as inadequate for today's global, interdisciplinary, and digital environment (Blass et al., 2012). Guerin et al. (2015) propose a collaborative, situated model that fosters reflexivity, resilience, creativity, and ethical judgement—qualities that are essential in an AI-rich knowledge economy.

Key dimensions of doctoralness include originality, criticality and reflexivity, methodological sophistication, scholarly communication (Kamler & Thomson, 2014), ethics and integrity (UNESCO, 2023), and scholarly identity within disciplinary communities (Wisker, 2012). Doctoral programmes should facilitate knowledge creation (McKenna, 2025). Moreover, doctoral education should cultivate autonomous researchers who possess the attributes expected of individuals capable of contributing to knowledge—what can be termed the cultivation of the 'knower' (Boud & Lee, 2009). These dimensions are interwoven and mutually reinforcing. As subsequent sections will argue, while AI can support these processes, it cannot replace the essential human capacities; thus, sustaining doctoralness requires deliberate attention as digital tools proliferate.

### 1.2 The PhD in the era of AI

As computer power increases with technological advancements, AI becomes a viable tool, swiftly becoming an integral part of society (Haenlein & Kaplan, 2019). While Generative Artificial Intelligence (GenAI) presents new challenges for doctoral education, it also offers an opportunity to refocus doctoral programmes on their fundamental purposes: contributing to knowledge and developing critical researchers (McKenna, 2025). According to Calvino et al. (2025):

> *The rapid rise of generative AI has sparked discussions about its potentially transformative effects and whether the technology will bring significant benefits in the form of widespread productivity increases. Despite the early evidence, generative AI appears to exhibit the defining characteristics of general-purpose technologies (GPTs): i) pervasiveness, ii) continuous improvement over time, and iii) innovation spawning. While productivity gains may not materialise immediately, the evolution of earlier GPTs seems to provide encouraging signs that generative AI could lead to substantial improvements in productivity in the future, notably through the innovation-spawning channel. The full realisation of generative AI's productivity potential in the long term will depend on the implementation of relevant policies (p. 3).*

This citation places HEIs at the mercy of technological advancement. Kariyana et al. (2017) established that, beyond institutional mediation, the dispositions of the doctoral supervisor and the doctoral supervisee are the core determinants of the quality of a doctoral candidate.

AI technologies now pervade the PhD lifecycle. Discovery tools and generative search models accelerate literature mapping; coding assistants support data cleaning and analysis; LLM-based writing aids provide feedback on structure, clarity, and argumentation; and analytics dashboards promise personalised progress monitoring for supervisors and candidates (Zawacki-Richter et al., 2019; Holmes et al., 2019; OECD, 2024). In laboratories and fieldwork settings, computer vision and Natural Language Processing (NLP) systems unlock new forms of data, while translation models lower linguistic barriers to global scholarly participation.

Alongside these technical advances, new governance questions arise: attribution of authorship in AI-mediated writing, data protection and consent in training sets, and transparency of methods and provenance. Institutions are drafting policies and guidance to help candidates and supervisors navigate these questions (UNESCO, 2023; European University Association, 2023). Equity remains a core concern: access to capable hardware, paid tools, and reliable connectivity is uneven, particularly across the Global South (Dlamini & Ndzinisa, 2025). Nevertheless, clearly, concerns about 'catching' students who 'cheat' the system and the acknowledgement of the benefits of Generative AI are not the whole story. The implications of Generative AI for doctoral studies are extensive and multifaceted (McKenna, 2025, pp. 2-3). McKenna (2025) continues:

If grappled with fully, these implications raise fundamental questions about the nature of knowledge creation, the role of the researcher, and the meaning of original scholarship. If the doctorate is simply about producing a thesis good enough to pass muster in the examination process, and Generative AI can produce such a thesis almost instantly, one might ask whether doctoral programmes should simply focus on training students in prompt engineering and ensuring they know how to scrutinise outputs for hallucinations and inaccuracies. This approach would certainly be more efficient, aligning with current imperatives for streamlined, cost-effective education delivery.

However, this would of course fail to address the purposes of doctoral education. The doctorate is not about knowledge creation in the mechanical sense of reviewing the literature and implementing the data collection and analysis methods accepted in the field; it is also fundamentally about nurturing the independent, responsible researcher who undertakes the work of critically engaging with literature and grappling with data in order to make a contribution (p.3).

The same sentiments are shared by many authors; however, what remains is the feasibility of promoting the holistic development of a doctoral student who exhibits the uncompromising graduate attributes expected at this level of education.

### 1.3 Problem statement

Quality postgraduate supervision is at the heart of postgraduate student success. To enhance supervisory experiences, various models have been developed over time. Recently, Iatrellis et al. (2025) argue that universities should consider incorporating AI tools, such as ChatGPT, to support PhD supervision, particularly in providing structured feedback and guidance. Supervisors should explore AI-assisted mentoring to optimise time-intensive advisory tasks and enhance research productivity (Iatrellis et al., 2025). However, despite the uptake of AI tools in doctoral research and supervision, doctoral education often lacks clear, discipline-sensitive frameworks for distinguishing productive AI assistance from practices that may erode the essence of doctoral work, particularly in under-resourced Global South contexts. The existing fragmented literature on the utilisation of AI in postgraduate supervision presents inconsistencies concerning the potential adoption of AI in research and supervision within the higher education sector. This gap creates uncertainty regarding the acceptable use of AI, how to evidence originality and independent scholarly judgement, and how to ensure research integrity and equity in AI-mediated research workflows. As this impacts candidates, supervisors, and institutions, this study aims to design supervision and assessment practices that safeguard the essence of doctoral work while enabling responsible innovation.

#### *1.3.1 Research objective*

To develop a framework that mediates the conceptualisation of AI-enabled practices influencing doctoralness and strategies for safeguarding doctoral-level graduate attributes in the digital age.

## 2. Methodology

This chapter adopts a qualitative, conceptual research design. It employs an integrative review of scholarly and policy literature to synthesise how generative AI intersects with doctoralness and doctoral supervision.

### 2.1 Data sources and search strategy

Sources comprised (i) peer-reviewed journal articles and book chapters on doctoral education/supervision and AI/GenAI, and (ii) key policy and guidance texts that shape doctoral education and responsible AI use (e.g., CHE standards, national reports, UNESCO guidance, and sectoral position statements). Searches were conducted using academic databases (Scopus, ERIC, and Google Scholar) and discovery tools. Search strings combined terms such as “doctoralness” OR “doctorateness” OR “doctoral education” OR “PhD supervision” with “artificial intelligence” OR “generative AI” OR “large language model” OR “ChatGPT” OR “academic integrity” OR “provenance.” Given the recency of GenAI, the search emphasised work published from 2019 to 2025 while also including earlier foundational scholarship on doctoral education and supervision for conceptual grounding. Reference lists of key sources were snowballed to locate additional relevant studies.

## 2.2 Inclusion and screening

Following the 2020 PRISMA flow diagram (Figure 1), records were screened for relevance to the chapter's objective through title/abstract review, followed by full-text screening. Items were included if they (a) addressed doctoral education, doctoral supervision, assessment of doctoral work, or doctoral graduate attributes, and (b) discussed AI/GenAI or closely related digital research tools with clear implications for doctoral formation and research practice. Literature with explicit relevance to South Africa or other Global South contexts was actively sought to support contextual sensitivity. Conceptual commentaries and policy statements were included selectively where they provided influential arguments or governance guidance directly relevant to the chapter. Ultimately, 31 papers were selected for this chapter.

## 2.3 Analysis and framework development

The retained sources were analysed using iterative qualitative content analysis. Extracted passages were coded for: (i) AI affordances across the doctoral lifecycle (literature mapping, analysis support, drafting, and revision), (ii) risks to doctoralness (integrity breaches, dependency, inequity, and opacity), and (iii) supervisory and assessment responses (disclosure, provenance records, feedback design, viva interrogation, and policy guidance). These themes were then mapped onto the CHE Knowledge and Skills graduate attributes (Table 1) to identify where AI can credibly support attribute development and where it threatens the evidencing of those attributes. The conceptual framework (Figure 2) was developed by iteratively refining the relationships between AI affordances and risks, mediating conditions, and doctoralness outcomes.

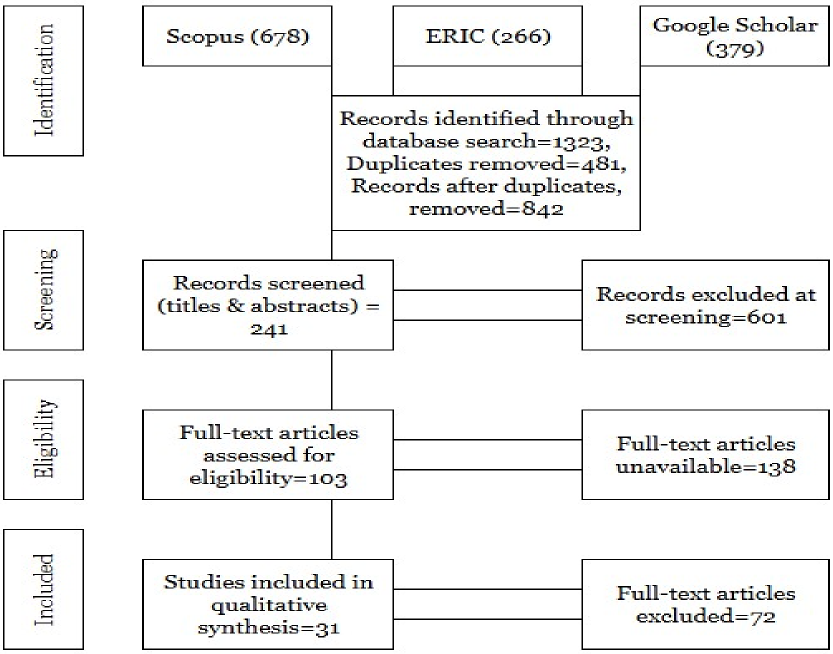



***Figure 1:*** *PRISMA 2020 flow diagram for the paper*

### 3.4 Ethical considerations

This chapter is desk-based and did not involve human participants or identifiable personal data. Therefore, ethical considerations centred on responsible scholarship, including the accurate representation of sources, careful handling of claims regarding integrity and misconduct, and a clear distinction between evidence-based findings and interpretive arguments.

## 4. Merits, Pitfalls and Supervision Dynamics in the Age of AI

The arrival of AI has rendered gains and losses to doctoral training.

### 4.1 Merits of doctoral training in the era of AI

The robustness of higher education institutions (HEIs) is inextricably linked to the quality of high school graduates, and their proficiency is largely dictated by the teaching methods they encounter (Tachie & Kariyana, 2022). To extend this narrative, we acknowledge that the quality of doctoral education, the pinnacle of education, is fundamental in shaping society. In this context, supervision experiences become essential. With GenAI, Akbar (2025) notes that the increasing use of AI tools in higher education necessitates a clear understanding of how students—by demography and study discipline—employ these tools, for what purposes, and how their use is evaluated. This understanding will inform the development of future guidelines and training for staff and students (Akbar, 2025).

Effective doctoral supervision is heavily dependent on dialogic feedback (Jensen et al., 2025). When we position the development of responsible, independent, and critical researchers at the heart of doctoral education, this has significant implications for how we engage with Generative AI in doctoral programmes (McKenna, 2025). The benefits of utilising AI include aiding research—for example, assisting and improving coding/programming, proofreading, writing, and serving as an explanatory tool for rendering complex information accessible (Akbar, 2025). Oliinyk et al. (2024) summarise that the ability of AI to optimise the work of future scientists, scholars, and academics on the topics of their research can be considered the main advantage.

When used judiciously, AI can enhance doctoral learning. First, it can create time and cognitive bandwidth by automating repetitive tasks—such as formatting, preliminary coding, and reference matching—allowing candidates to invest more effort in theory building and interpretation (Kumar & Gunn, 2025). Second, AI can broaden access by assisting multilingual writers and neurodiverse learners, thereby supporting inclusive pedagogy and epistemic justice (UNESCO, 2023). Third, tools that surface diverse literatures can spur interdisciplinarity and creativity, while simulation and modelling expand methodological repertoires (Barrett & Pack, 2023). Finally, AI-enabled collaboration platforms and open-science workflows can connect dispersed teams and shorten the path from idea to impact.

These benefits depend on explicit human oversight and reflective practice: AI suggestions must be interrogated, triangulated with evidence, and situated within disciplinary norms—activities that themselves deepen the doctoral experience.

### 4.2 Pitfalls of doctoralness in the age of AI

The increasing prevalence of Artificial Intelligence (AI) in higher education underscores the necessity to explore its implications for ethical, social, and educational dynamics within the sector (Al-Zahrani & Alasmari, 2024). Risks cluster around integrity, dependency, inequity, and opacity. Generative systems challenge conventions of authorship and novelty, raising questions about what constitutes an original contribution (Thorp, 2023; Nguyen et al., 2024). Over-reliance can erode critical thinking and methodological understanding if candidates outsource judgment to opaque models (Yan et al., 2024). Biases embedded in training data may reproduce epistemic injustices, while uneven access to premium tools risks widening gaps between well-resourced and under-resourced institutions (Dlamini & Ndzinisa, 2025). Supervisors and examiners also face difficulties in verifying provenance and ensuring transparent reporting of AI involvement (Cotton et al., 2024; Kofinas et al., 2024).

The quality and validity of the information AI produces vary based on the AI product and the logistics of implementation (Mueller & Massaron, 2022). Factors affecting the quality of a doctoral graduate are embedded in the characteristics of universities and doctoral students (Kariyana et al., 2017). In confirmation, Ungadi's (2021) study found that multiple issues coalescing into challenges stemmed from PhD students' and PhD supervisors' past experiences, the structures in place to facilitate doctoral education, and the intersection between these structures, PhD supervisors, and the context of doctoral studies. Boyd and Harding (2025) highlight that the often-unacknowledged use of GenAI in doctoral research can confer undue agency on the technology, disrupting traditional relationships in an unacknowledged manner. The rapid but often unacknowledged uptake of GenAI within doctoral research occurs alongside a lack of consideration for the emotional support students attribute to the technology.

Potential problems include environmental costs, violations of intellectual property rights, the provision of misleading and/or inaccurate information, and risks of plagiarism and hindered creativity (Akbar, 2025). Opportunities to develop critical analytical skills are missed, which can threaten the integrity of research outputs (Harding & Boyd, 2023). Oliinyk et al. (2024) state that the challenge lies in the violation of ethics and academic integrity by graduate students (p. 302). Doctoral supervisors face numerous challenges in overseeing doctoral students. Such challenges include the structured nature of universities, which requires most experienced, skilled, and knowledgeable academic staff to engage in multiple administrative responsibilities, increasing their workload, and reducing the time available for PhD supervision. Additionally, contemporary doctoral candidates often require substantial support due to their weaker academic abilities (Ungadi, 2021).

### 4.3 Supervision dynamics in the age of AI

There is no doubt that the concept of supervision has evolved over time and will continue to do so. This evolution is evident in the transitions in postgraduate supervision models, which reflect a dissatisfaction with the quality of supervision over time and across different contexts, and a call for improvement. Throughout history, there has been no episode that has posed a sudden threat to supervisory roles, except for AI. For this reason, it is sensible to dedicate a section of this paper to positioning ourselves in relation to this reality. The Council on Higher Education (CHE) Qualification Standard for Doctoral Degrees (NQF Level 10) serves as the organising analytical lens.

It is inevitable that a balance must be struck between competitors if the market is to attain equilibrium. In this case, the long-held human supervision relationship seems to be disrupted by AI, which is the new kid on the block. Suddenly, it appears that supervisors have lost grip on their supervisees (Harding & Boyd, 2023). The advent and adoption of GenAI tools have pedagogical implications for researcher/supervisor dynamics. The rise of generative artificial intelligence chatbots raises the question of how interactions with a chatbot align with, or diverge from, authentic feedback practices with supervisors (Jensen et al., 2025).

ChatGPT can enhance PhD supervision by providing structured academic recommendations, reducing administrative burdens on supervisors, and contributing to the evolution of a "tripartite mentoring model" where AI, supervisors, and students collaborate to tackle complex research challenges (Iatrellis et al., 2025). The present study suggests that prospective students would benefit from being informed about the limitations of generative AI, as well as the risks it poses to accuracy, originality, creativity, and fostering dependence. It would also be prudent to outline the threat that training AI models poses to intellectual property rights, as under current trends, doctoral students' future work is likely to be utilised by such models without their consent (Akbar, 2025).

The supervisory role is evolving from sole gatekeeper to AI-literate mentor. Supervisors need fluency in the capabilities and limitations of current tools, as well as strategies for fostering students' reflective use of them (Caillaud & Skec, 2024; Wisker, 2012). Practical steps include co-creating AI-use agreements, requiring reflective memos that justify when and how AI was used, and designing assessments that privilege process, reasoning, and methodological justification over surface polish. Supervisors should model scholarly integrity by disclosing their own AI use and prioritising formative dialogue about argument quality and evidence.

Findings highlight that fear and suspicion surrounding the use of GenAI confer undue agency on the technology, which further conceals its use (Harding & Boyd, 2023).Chatbot feedback encounters highlighted the student's agency and focused on the task; supervisor feedback encounters were relational, contextual, and developmental (Jensen et al., 2025). AI-generated recommendations were most effective when structured around topic-specific concepts (Iatrellis et al., 2025). The study established that the majority of PhD students have higher levels of research competence in terms of substantive and design components, and lower levels in terms of procedural, optional, and communicative components. The results of the study indicate that AI use is appropriate for increasing the research competence of future scientists, scholars, and academicians (Oliinyk et al., 2024).

## 5. Doctoralness in South Africa

### 5.1 Council on higher education graduate attributes for doctoral degrees

The Council on Higher Education’s Qualification Standard for Doctoral Degrees defines 'doctoralness' through two interlinked sets of graduate attributes—Knowledge and Skills—that are used nationally as the threshold for awarding doctoral qualifications (National Qualifications Framework [NQF] Level 10) (Council on Higher Education, 2018; Boughey, 2023). In South Africa, the Higher Education Qualifications Sub-Framework describes this as making ‘a significant and original academic contribution at the frontiers of a discipline or field’ (CHE, 2013).

Knowledge attributes include: broad and current field knowledge; expert, in-depth knowledge in the specific research area; insight into connections with cognate fields; ethical awareness in research and professional conduct; and evidence of an original contribution that advances the field. Skills attributes require candidates to: evaluate, select and apply appropriate research approaches and methods; work with reflection and autonomy to reach defensible conclusions; communicate effectively (including information and digital literacy); and demonstrate critical and analytical thinking. These attributes provide the anchor points against which AI’s role should be judged in this chapter.

### 5.2 Implications from the CHE doctoral degrees national report (2022)

The CHE’s Doctoral Degrees National Report applies the same attribute framework to evaluate provision nationally. It highlights three implications relevant to AI-rich doctoral education: (1) Make attributes explicit and monitor progression throughout candidature (not only infer them at the thesis stage); (2) Strengthen supervision—ensure supervisors intentionally cultivate autonomy, ethical awareness, and critical judgement, including through AI-use agreements and provenance statements; (3) Balance disciplinary depth with cognate awareness—use AI to surface cross-field links while maintaining methodological fit and epistemic integrity (Leitch et al., 2022). In short, align institutional AI policies, supervision practices, and assessment rubrics directly with the CHE Knowledge and Skills attributes.

*Table 1: Mapping CHE doctoral graduate attributes to AI-use guardrails*

| Attribute domain | CHE attribute (short label) | How AI can help | Risks to doctoralness | Concrete guardrails |
|---|---|---|---|---|
| Knowledge | Broad and current field knowledge | Landscape scans; topic overviews; rapid alerts | Uncritical summaries; superficial breadth | Annotated bibliographies; primary-source verification; log prompts & sources |
| Knowledge | Expert, in-depth knowledge in area | Corpus building; code/data assistance | Methodological shortcuts; dependency on models | Supervisor-approved methods or plan; reproducible notebooks; viva on design rationale |
| Knowledge | Interconnectedness with cognate fields | Cross-disciplinary retrieval; mapping related work | Hallucinated links; scope drift | Justify inclusion criteria; provenance records; scoped review protocol |
| Knowledge | Ethical awareness and conduct | Policy checks; consent templates; data-risk flags | Privacy/IP breaches; opaque model use | Provenance statements; data governance review; ethics sign-off including AI section |
| Knowledge | Original contribution | Gap analysis; simulation scaffolds | Plagiarism; derivative contributions | Originality statement; similarity checks; oral defense on novelty |
| Skills | Select and apply appropriate methods | Method comparison; diagnostics | Method mismatch; overfitting via automation | Methods-justification rubric; (pre)registration where applicable |
| Skills | Reflection and autonomy | Self-quiz; revision planners | Over-delegation to tools | Weekly reflection logs; supervisor attestation of independent work |
| Skills | Scholarly communication (including digital literacy) | Language polishing; visualisation assistance | Loss of authorial voice; hidden edits | Editing disclosure; keep tracked changes; retain pre-AI drafts |
| Skills | Critical and analytical thinking | Counterargument generation; error-finding | Shallow critique; confirmation bias | Structured critique templates; require manual replication/ablation where relevant |

## 6. Safeguarding and enriching doctoralness, and looking beyond

Safeguarding doctoralness requires institutional frameworks that enable responsible innovation. Core elements include: (1) clear, discipline-sensitive policies for transparent AI use and provenance; (2) curricular AI literacy for candidates and supervisors; (3) assessment practices that evaluate higher-order reasoning and originality; and (4) infrastructure and procurement choices that reduce inequities (UNESCO, 2023; European University Association, 2023; Leitch et al., 2022). In the Global South, partnerships that share computing resources and training, as well as the prioritisation of open-source tools, can help mitigate access gaps (Akala, 2021). Embedding reflexivity and epistemic justice perspectives ensures that AI augments, rather than distorts, disciplinary knowledge-making.

Throughout the doctoral journey, candidates are expected to display doctorateness in their thesis via the characteristics of high-quality scholarly research (Trafford & Leshem, 2009). Competencies expected of doctoral graduates include being autonomous researchers and knowledge producers (Kariyana et al., 2017). The presence of chatbots underlines, rather than replaces, the need for doctoral supervisors' engagement with feedback practices that lead to meaningful learning (Jensen et al., 2025). Unless both parties understand these implications, GenAI tools have the potential to disrupt the traditional balance of power and trust between Researcher and Supervisor, potentially impacting both the rigour of PhD training and research outcomes (Harding & Boyd, 2023). Perhaps most immediately, doctoral candidates need to engage in sustained reflection about what it means to work with AI and what it means to allow AI to perform significant portions of their scholarly work. Have they contributed to knowledge if AI has generated significant portions of the analysis or writing? What kind of critical researchers are they becoming through this process? These questions go to the heart of doctoral education's purposes (McKenna, 2025).

### 6.1 Future directions

Looking ahead, doctoral education is likely to incorporate artificial intelligence (AI) more profoundly into tailored learning pathways, project management, and collaborative knowledge creation. Scenario analysis indicates an increasing utilisation of intelligent research environments that monitor provenance, support reproducibility, and offer formative analytics to learners and supervisors (Oliinyk et al., 2024; OECD, 2024). Global policy frameworks (UNESCO, 2023; European University Association, 2023) will shape standards for disclosure, authorship, and data governance. Importantly, conceptions of originality may evolve from the production of text to the generation of defensible reasoning, designs, and contributions to open, machine-actionable knowledge bases.

The integration of Generative AI into doctoral education thus presents both challenges and opportunities. By confronting these challenges directly, through critical education regarding AI's capabilities and limitations, ongoing reflection on ethical implications, and a renewed emphasis on scholarly responsibility, we can leverage this moment to strengthen rather than undermine the fundamental purposes of doctoral education (McKenna, 2025). The study concludes that Generative AI tools should be more than a covert "third wheel" in the relationship. Instead, the technology could be openly incorporated into supervision frameworks in a transparent and integrated manner (Harding & Boyd, 2023; Boyd & Harding, 2025).

## 7. Parting Shot

Artificial intelligence is not a surrogate for doctoralness; at best, it is a demanding collaborator. Throughout this chapter, we have treated the CHE Qualification Standard for Doctoral Degrees (2018) as the anchor for what counts at NQF Level 10, examining contemporary AI through that lens. The Standard's Knowledge and Skills attributes delineate the destination; AI is merely a set of vehicles that may speed or skew the journey. Our analysis shows that when institutions make those attributes explicit, teachable, assessable, and auditable, AI can be integrated without hollowing out the core of doctoral formation.McKenna (2025) concluded that:

> *In the South African context, where transformation remains a central imperative for higher education, refocusing on the substantive purposes of doctoral education—namely, knowledge creation that serves society and the development of ethically responsible scholars—directly addresses transformation goals. By foregrounding questions about whose knowledge counts, whom that knowledge serves, and what responsibilities accompany scholarly privilege, we create conditions for doctoral education that can genuinely contribute to social justice and decolonial knowledge production (p.5).*

### 7.1 Conceptual framework

In this regard, we propose a mediating conceptual framework. Figure 1 presents the proposed conceptual framework. The framework positions artificial intelligence (AI) as a set of research and supervision capabilities (affordances) that can either support or threaten doctoralness, depending on mediating conditions that function as guardrails. AI affordances (e.g., accelerated literature mapping, coding support, drafting/translation, and feedback/progress analytics) can enhance efficiency and access. As Iatrellis et al. (2025) found, ChatGPT demonstrated the capacity to refine research methodologies and improve knowledge discovery.

Conversely, AI risks (e.g., hallucinations, plagiarism/ghost authorship, bias, privacy/intellectual property breaches, and dependency) can undermine originality, critical judgement, and integrity. The net effect on doctoralness—anchored in the CHE NQF Level 10 Knowledge and Skills attributes—is mediated by (i) candidate AI literacy and reflexivity, (ii) supervisor AI pedagogy and feedback practices, (iii) institutional policy and ethics governance, (iv) access and infrastructure (equity), and (v) disciplinary norms and assessment design.

The framework, therefore, provides an analytical scaffold for connecting the literature reviewed in this chapter to actionable guidance for doctoral candidates, supervisors, and institutions.

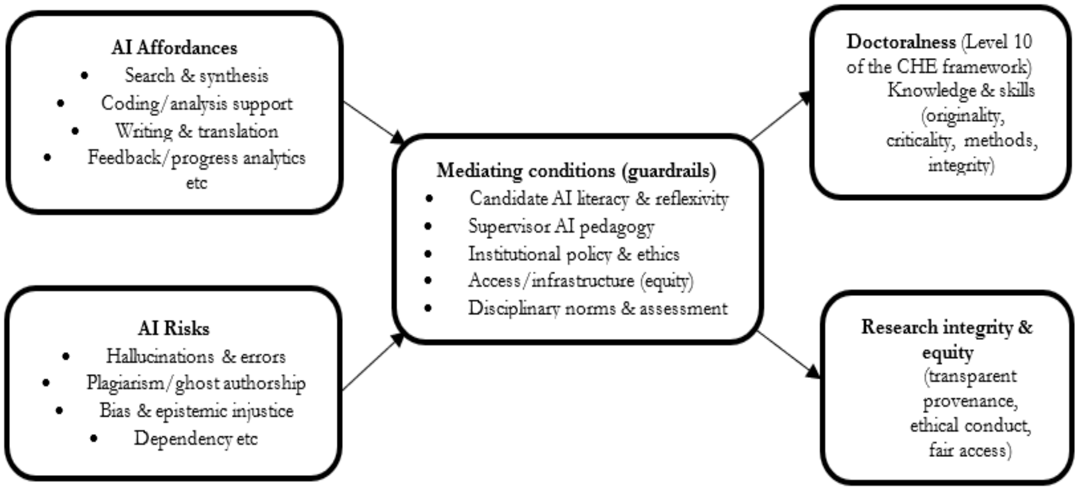



***Figure 1:*** *Conceptual framework for AI-enabled doctoral education and doctoralness*

Throughout this discussion, we emphasise that maintaining the essence of doctoral scholarship in a digital age requires not the rejection of artificial intelligence, but its critical and ethical appropriation – accompanied by ongoing monitoring and reflection – as a collaborator rather than a replacement in advanced research learning. In this chapter, we acknowledge certain limitations, including the non-exhaustive nature of an integrative review, the rapid evolution of generative AI tools and guidance (which may outpace publication cycles), and the uneven availability of evidence from under-resourced contexts. Consequently, the analysis seeks to achieve conceptual clarity and practical implications rather than definitive claims of prevalence.

## 8. Conclusion and Recommendations

First, the thesis alone constitutes an insufficient proxy for the graduate attributes. The Doctoral Degrees National Report (2022) elucidates that many programmes still infer attributes retrospectively at the point of submission. In an AI-rich environment, such an approach becomes untenable. Provenance, methodological judgement, and independent reasoning must be cultivated and evidenced throughout the candidature, rather than reconstructed post hoc. The practical implication is to scaffold each CHE attribute with observable behaviours and artefacts, including research design defences, critical appraisal diaries, methods justifications, and ethics addenda that encompass tool use.

Second, supervision represents the decisive site of assurance. The potential of AI—namely, rapid literature mapping, coding assistance, and diagnostic checks—can only be realised when supervisors intentionally design tasks that reward critical thinking rather than mere surface refinement. Routine text-polishing may be permitted with appropriate disclosure; however, interpretation, critique, design choices, and argumentation must remain the candidate's work and be assessed as such. We recommend programme-wide supervision workshops that translate the CHE attributes into assessment prompts (for example, "defend your choice of estimator against viable alternatives, including those suggested by an AI assistant") and oral examinations that explore tool-mediated processes.

Third, assessment regimes necessitate a modest but consequential redesign. In an environment where AI threatens to commodify superficial outputs, assessment must pivot towards process evidence: versioned research notebooks, preregistered design rationales (where appropriate), and viva voce components focusing on decisions that AI cannot credibly claim (such as assumption checks, trade-offs, and ethical reasoning). Assessment rubrics should explicitly allocate marks to original contributions, as demonstrated by novelty claims, triangulated gap analyses, and examiner-tested argument structures, rather than to textual fluency or diagrammatic presentation.

Fourth, provenance and ethics must be normalised as integral components of scholarly practice. A lightweight, programme-standard AI provenance statement—detailing tools, versions, prompts/workflows, and human verification steps—should be appended to chapters, articles, code, and datasets. Ethics review forms should similarly pose clear questions regarding data flows, privacy, intellectual property, and biases inherent in models and training sets. These practices are not bureaucratic add-ons; they are essential mechanisms through which integrity and accountability, as articulated within the CHE attributes, are rendered visible.

Fifth, equity and capability warrant sustained attention. Artificial Intelligence (AI) has the potential to both exacerbate and mitigate inequalities. Candidates with limited computational resources, connectivity, or institutional subscriptions may be directed towards inferior tools; conversely, an over-reliance on AI could undermine the very skills that the Standard mandates. Consequently, programmes should integrate access strategies (such as shared infrastructure and curated open-source pathways) with capability-building initiatives (including critical AI literacy, error identification, and methodological judgement). The aim is not universal adoption of AI but rather discerning utilisation that aligns with disciplinary norms and the expectations set forth in the Standard.

Sixth, institutions require cohesive policy that is succinct, enforceable, and pedagogically transparent: (1) what is permissible concerning disclosure; (2) what is prohibited; (3) what must always be evidenced; and (4) how breaches relate to existing misconduct frameworks. Importantly, policy should be co-created with examiners and supervisors to ensure that it translates seamlessly into tasks, rubrics, and viva protocols, thereby closing the loop between doctrine and practice.

Finally, we advocate for a targeted research and improvement agenda: (a) to develop validated instruments for measuring progression on each CHE attribute under conditions enriched by AI; (b) to trial viva formats that emphasise "explain-your-decision" questioning of tool-assisted steps; (c) to compare learning designs that differentially prioritise process artefacts versus final products; and (d) to investigate supervisor development models that most effectively enact change in practice. Such endeavours should report not only on outcomes (completion rates, quality) but also on integrity metrics (such as provenance quality, success rates of replicating analysis pipelines, and examiner confidence in originality claims).

In summary, AI-enabled practices represent the 'affordances' aspect of the model and are examined in relation to the concept of doctoralness and the mediating frameworks. Preserving doctoralness in a digital age does not necessitate a ban on novel tools; rather, it calls for a re-centring of what the doctorate certifies and the back-designing of curricula, supervision, assessment, and policy to ensure that AI occupies its appropriate role: as a fallible assistant serving the candidate's judgement, creativity, and contribution. With the CHE Standard serving as the guiding principle and the lessons from the National Report functioning as practical benchmarks, South African programmes have the potential to leverage AI to enhance, rather than diminish, the significance of the PhD.

## 9. Declarations

**Funding:** This research did not receive any external funding.
**Conflicts of Interest:** The authors declare no conflict of interest.
**Use of Artificial Intelligence:** The current work was created with minimal assistance from artificial intelligence technologies, mainly Grammarly to refine language for clarity, and ChatGPT to conceptualise the development of Figure 1.

## References

Akala, B. U. (2021). Challenges in doctoral supervision in South African universities. *Open Science Journal, 6*(2), 1–20. https://doi.org/10.23954/osj.v6i2.2886

Akbar, M. N. (2025). Use of artificial intelligence tools by doctoral students: A mixed-methods explanatory-sequential investigation. *Journal of Further and Higher Education, 49*(7), 995–1013. https://doi.org/10.1080/0309877X.2025.2515135

Al-Zahrani, A.M., & Alasmari, T.M. (2024). Exploring the impact of artificial intelligence on higher education: The dynamics of ethical, social, and educational implications. *Humanities and Social Sciences Communications, 11*, 912. https://doi.org/10.1057/s41599-024-03432-4

Backhouse, J. P. (2009). *Doctoral education in South Africa: Models, pedagogies and student experiences* (Doctoral dissertation, University of the Witwatersrand, Johannesburg, South Africa).

Barrett, A., & Pack, A. (2023). Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education, 20*(1), 59.

Blass, E., Jasman, A., & Levy, R. (2012). Supervisor reflections on developing doctoralness in practice-based doctoral students. *Quality Assurance in Education, 20*(1), 31–41.

Boud, D., & Lee, A. (2009). *Changing practices of doctoral education*. Routledge.

Boughey, C. (2023). *The qualification standard for doctoral degrees and its implications for the pedagogy of doctoral supervision and assessment (Briefly Speaking No. 26)*. Pretoria: Council on Higher Education.

Boyd, P., & Harding, D. (2025). Generative AI: Reconfiguring supervision and doctoral research. *Buildings and Cities, 6*(1), 294–309. https://doi.org/10.5334/bc.560

Caillaud, E., & Skec, S. (2024, September). Supervision of design PhD students in an era of generative artificial intelligence. In the *26th International Conference on Engineering and Product Design Education* (pp. 420–425).

Calvino, F., Haerle, D., & Liu, S. (2025). Is generative AI a general-purpose technology? Implications for productivity and policy. *OECD Artificial Intelligence Papers: June 2025 No. 40*. Paris: OECD.

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239.

Council on Higher Education (CHE). (2018). *Higher education qualifications sub-framework qualification standard for doctoral degrees*. Pretoria: Council on Higher Education.

Council on Higher Education (CHE). (2018). *Qualification Standard for Doctoral Degrees*. Pretoria: CHE.

Dlamini, R., & Ndzinisa, N. (2025). Towards a critical discourse on artificial intelligence and its misalignment in sub-Saharan Africa: Through an equality, equity, and decoloniality lens. *Journal of Education, 98*, 42–61.

European University Association (EUA). (2023, February 14). *Artificial intelligence tools and their responsible use in higher education learning and teaching.* https://www.eua.eu/publications/positions/artificial-intelligence-tools-and-their-responsible-use-in-higher-education-learning-and-teaching.html

Guerin, C., Kerr, H., & Green, I. (2015). Supervision pedagogies: Narratives from the field. *Teaching in Higher Education, 20*(1), 107–118. https://doi.org/10.1080/13562517.2014.957271

Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review, 61*(4), 5–14. https://doi.org/10.1177/0008125619864925

Harding, D., & Boyd, P. (2023). Generative AI and PhD supervision: A covert third wheel or a seat at the table? *University College of Estate Management (UCEM) Institutional Research Repository*.https://ube.repository.guildhe.ac.uk/id/eprint/125/

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning.* Center for Curriculum Redesign.

Iatrellis, O., Bania, A., Samaras, N., Kosmopoulou, I., & Panagiotakopoulos, T. (2025). ChatGPT in doctoral supervision: Proposing a tripartite mentoring model for AI-assisted academic guidance. *International Journal of Doctoral Studies*, *20*, Article 9. https://doi.org/10.28945/5579

Kamler, B., & Thomson, P. (2014). *Helping doctoral students write: Pedagogies for supervision.* Routledge.

Kariyana, I., Sonn, R. A., & Marongwe, N. (2017). Objectivity of the subjective quality: Convergence on competencies expected of doctoral graduates. *Cogent Education*, *4*(1), 1-15. https://doi.org/10.1080/2331186X.2017.1390827

Kofinas, A., Bentley, Y., O'Keeffe, C., & Patrick, K. (2024). The impact of generative AI on academic integrity of authentic assessments within a higher education context. *British Journal of Educational Technology.* (Advance online publication).

Kumar, S., & Gunn, A. (2025). Doctoral students' reflections on generative artificial intelligence (GenAI) use in the literature review process. *Innovations in Education and Teaching International*, *62*(4), 1395–1408.

Jensen, L. X., Bearman, M., Boud, D., & Konradsen, F. (2025). Feedback encounters in doctoral supervision: The role of generative AI chatbots. *Assessment & Evaluation in Higher Education,*https://doi.org/10.1080/02602938.2025.2478155

Leitch, A., Burton, S., Ntshoe, I., Kaniki, A., & Ralebipi-Simela, R. (2022). *National review of South African doctoral qualifications 2020–2021: Doctoral degrees national report.* Pretoria: Council on Higher Education.

Maurya, R. K. (2023). Using AI-based chatbot ChatGPT for
practicing counseling skills through role-play. *Journal of Creativity in Mental Health,19*(4), 513–528. https://doi.org/10.1080/15401383.2023.2297857

McKenna, S., & van Schalkwyk, S. (2024). A scoping review of the changing landscape of doctoral education. *Compare: A Journal of Comparative and International Education,54*(6), 984–1001.

McKenna, S. (2025). Foregrounding doctoral knowledge and knower in the age of generative artificial intelligence. *Transformation in Higher Education,10*(0), a653. https://doi.org/10.4102/the.v10i0.653

Mowbray, S., & Halse, C. (2010). The purpose of the PhD: Theorising knowledge as work. *Higher Education Research & Development,29*(6), 653–664. https://doi.org/10.1080/07294360.2010.487199

Mueller, J. P., & Massaron, L. (2022). *Artificial Intelligence for Dummies.* John Wiley & Sons, Inc.

Nguyen, A. K. T., Kobayashi, M., & Hohlfeld, T. N. (2024). Human–AI collaboration patterns in AI-assisted academic writing in higher education. *Studies in Higher Education,49*(5). https://doi.org/10.1080/03075079.2024.2323593

OECD. (2024). *Artificial intelligence and education and skills.* Organisation for Economic Co-operation and Development. https://www.oecd.org/en/topics/sub-issues/artificial-intelligence-and-education-and-skills.html

Oliinyk, I., Bulavina, O., Romanenko, T., Tatarnikova, A., & Smirnov, A. (2024). Artificial intelligence in developing doctoral students' research competencies. *Revista Eduweb,18*(3), 294-305. https://doi.org/10.46502/issn.1856-7576/2024.18.03.22

Sarrico, C. S. (2022). The expansion of doctoral education and the changing nature and purpose of the doctorate. *Higher Education,84*(6), 1299–1315.

Tachie, S. A., & Kariyana, I. (2022). Time divested or time invested? Freshmen's perspectives and reflective experiences on interactive school learning. *Research in Social Sciences and Technology,7*(1), 1-19. https://doi.org/10.46303/ressat.2022.2

Thorp, H. H. (2023). ChatGPT is fun, but not an author. *Science,379*(6630), 313.

Trafford, V., & Leshem, S. (2009). Doctorateness as a threshold concept. *Innovations in Education and Teaching International, 46*(3), 305–316. https://doi.org/10.1080/14703297.2009.0307015

UNESCO. (2023, September 7). *Guidance for generative AI in education and research.*https://unesdoc.unesco.org/ark:/48223/pf0000386693

Ungadi, B. A. (2021). Challenges in doctoral supervision in South African universities. *Open Science Journal, 6*(2), 1–24. https://doi.org/10.12691/osj-6-2-1

Wisker, G. (2012). *The good supervisor: Supervising postgraduate and undergraduate research for doctoral theses and dissertations.* Bloomsbury Publishing.

Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology, 55*(1), 90–112. https://doi.org/10.1111/bjet.13369

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education, 16,* 39. https://doi.org/10.1186/s41239-019-0171-0

CHAPTER FIVE

# Exploring Supervisors' Readiness to Integrate AI Tools in Postgraduate Supervision in Higher Education

**Maryam C. Adewunmi**[1]

AFFILIATIONS

[1]Centre for Education Rights and Transformation, Faculty of Education, University of Johannesburg, Johannesburg, South Africa.

CORRESPONDENCE

Email: hi2chinasa@gmail.com*

REFERENCE

Adewunmi, M. C. (2026). Exploring supervisors' readiness to integrate AI tools in postgraduate supervision in higher education. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 74–92). ERRCD Forum. https://doi.org/10.38140/obp4-2026-05



**Abstract:** The integration of artificial intelligence (AI) tools in postgraduate supervision has the potential to transform research processes by enhancing the quality of feedback, improving the efficiency of supervision, and addressing persistent challenges such as time constraints and student engagement within higher education. Despite this promise, the readiness of supervisors to adopt AI remains uneven, necessitating an exploration of their preparedness to integrate such tools. This study employs a constructivist paradigm and a qualitative research approach, guided by a generic qualitative research design. It draws on semi-structured interviews with 15 postgraduate supervisors from diverse disciplines to examine their perspectives on the integration of AI tools in postgraduate supervision. Through thematic analysis, four central themes emerged: technological literacy, institutional support, perceptions of AI, and ethical considerations, revealing the complex interplay between individual competence and institutional context. Supervisors with prior experience in digital technologies or from technology-intensive fields demonstrated higher readiness, while those from non-technical backgrounds encountered challenges due to limited digital exposure and perceived complexity. Institutional factors such as digital infrastructure, supportive policies, and professional development opportunities further influenced readiness levels. However, concerns surrounding academic rigour, ethical accountability, and workload pressures continue to constrain adoption. The chapter concludes by recommending targeted capacity-building programmes, institutional policy reforms, interdisciplinary collaboration, and enhanced supervisor–student partnerships to ensure ethical and effective AI use. Ultimately, while AI tools hold significant potential to enhance supervision efficiency and personalised support, their successful implementation requires tailored strategies responsive to diverse supervisory contexts, offering valuable insights for higher education institutions seeking to promote responsible AI integration in postgraduate supervision.

***Keywords:*** AI tools, postgraduate supervision, readiness, supervisor-student collaboration, technological literacy.

## 1. Introduction

The emergence of Artificial Intelligence (AI) has precipitated significant transformations across various sectors, including higher education. In recent years, the adoption of AI tools has become increasingly prevalent, aimed at enhancing administrative processes, teaching, learning, and research (Lee et al., 2024; Tan et al., 2024). Within the realm of postgraduate education, AI technologies such as ChatGPT, Grammarly, and research-based analytics tools present potential benefits for the improvement of quality and efficiency in supervision (Nouri et al., 2020). These tools can aid in generating feedback, identifying academic sources, checking for plagiarism, managing data, and providing writing support—functions traditionally performed manually by supervisors (Balalle & Pannilage, 2025).

Despite this promise, the preparedness of academic supervisors to integrate AI tools into postgraduate supervision is an area that remains underexplored, particularly in developing countries. Recent studies indicate that while the adoption of AI is advancing rapidly in the domains of teaching and assessment, the supervisory aspect is significantly under-researched, especially in low- and middle-income contexts where digital transformation is uneven (Omodan, 2025). Evidence suggests that supervisors' readiness is frequently impeded by structural inequities, including limited technological infrastructure, variable digital literacy, and inconsistent institutional strategies for AI integration (Akgun & Greenhow, 2022; Bayly-Castaneda et al., 2024). These deficiencies underscore the necessity for targeted empirical inquiry into supervisors' preparedness, given that their role is crucial in shaping the quality of postgraduate research and promoting ethical engagement with AI.

Recent scholarship confirms that empirical research on supervisors' readiness for AI integration is scarce, with the majority of studies focusing instead on student usage patterns or institutional digital strategies, rather than the readiness of supervisors (Chan & Hu, 2023; Nguyen et al., 2024). Research conducted in developing contexts further illustrates a widening gap in knowledge and skills, highlighting that supervisors often possess limited technological exposure and inadequate institutional support necessary for the effective adoption of AI tools (Kassa & Worku, 2025). This gap in evidence accentuates the significance of examining supervisory readiness, especially within resource-constrained higher education systems where digital transformation is inconsistent. Readiness comprises the knowledge, skills, attitudes, and institutional support structures vital for successful adoption (Uren & Edwards, 2023). As the responsibilities of supervisors evolve with technological advancements, it is imperative to understand their preparedness to engage with AI in ways that uphold academic integrity and facilitate student success (Rasul et al., 2024). This study aims to explore supervisors' readiness to integrate AI tools in postgraduate supervision within the broader context of digital transformation in higher education.

While the integration of artificial intelligence (AI) is progressively transforming the educational landscape, empirical evidence regarding the preparedness of postgraduate supervisors to incorporate AI tools into their supervision practices remains limited (Omodan, 2025). Many institutions are either unaware of or ill-prepared to address the pedagogical and ethical implications associated with AI usage in postgraduate research (Lee et al., 2024). Supervisors may lack the necessary technical skills, institutional support, or confidence to effectively utilise AI tools, thereby impeding innovation and adversely affecting the quality of research supervision (Kassa & Worku, 2025). Furthermore, in the absence of a comprehensive understanding of supervisors' readiness, higher education institutions may encounter difficulties in designing appropriate training or support systems to facilitate effective integration. This study aims to address this gap by investigating the extent of supervisors' readiness to incorporate AI tools in postgraduate supervision. Although existing research has underscored the potential of AI to enhance teaching and learning within higher education (Kassa & Worku, 2025; Nouri et al., 2020), a significant gap persists in studies specifically focusing on the preparedness of academic staff, particularly supervisors, to integrate AI into postgraduate supervision (Sehmi et al., 2025). Most extant literature predominantly emphasises student utilisation of AI, ethical concerns, or general perceptions of AI in education (Chan & Hu, 2023; Nguyen et al., 2023; Rasul et al., 2024; Uren et al., 2023), with insufficient attention directed towards supervisors' skills, attitudes, and preparedness. Scholars increasingly caution that this omission is problematic, as supervisors play a critical role as gatekeepers of research quality, ethical conduct, and methodological rigour, rendering their readiness essential for responsible AI integration (Spring et al., 2022; Malik et al., 2023). In the absence of intentional capacity-building and empirically grounded insights into supervisors' needs, institutions risk exacerbating digital divides, reinforcing inequities across disciplines, and compromising supervision standards in the context of generative AI (Makore, 2024; Rioseco-Pais et al., 2024). Understanding supervisors' readiness to integrate AI tools into postgraduate supervision is vital for guiding institutional policy, staff development, and digital transformation strategies. As higher education institutions endeavour to uphold quality supervision and support research excellence in the digital age, it becomes imperative to assess and enhance supervisors' competencies related to AI usage.

### 1.1 Problem statement

The rapid expansion of artificial intelligence (AI) in higher education has significantly transformed teaching, learning, and research practices; however, its integration into postgraduate supervision remains uneven and under-examined. While AI tools present clear potential to enhance supervisory efficiency, feedback quality, and student support, existing research has predominantly focused on student utilisation and institutional strategies, with limited empirical attention directed towards supervisors as key agents of postgraduate research quality and ethical governance. This gap is noteworthy, as supervisors play a central role in shaping research rigour, academic integrity, and responsible engagement with emerging technologies. In the absence of a clear understanding of supervisors' readiness to integrate AI, institutions risk promoting AI adoption without adequately considering supervisors' skills, attitudes, ethical concerns, and disciplinary contexts.

This issue is particularly pronounced in developing and resource-constrained higher education contexts, where digital transformation is inconsistent and institutional support structures are often fragmented. Evidence from this study indicates that supervisors' readiness to integrate AI varies considerably across disciplines and is influenced by technological literacy, perceptions of AI, workload pressures, and the availability of institutional guidance and training. Without coherent policies and targeted capacity-building, AI integration in postgraduate supervision risks becoming ad hoc, reinforcing disciplinary and digital inequalities and potentially compromising supervision quality and academic integrity. There is, therefore, a pressing need for empirically grounded insights into supervisors' readiness to inform responsible, equitable, and contextually responsive AI integration in postgraduate supervision. Hence, the study answer the following question: *What is the supervisors' readiness to integrate AI tools into postgraduate supervision in higher education?*

## 2. Methodology

This study is grounded in the constructivist research paradigm, which posits that reality is socially constructed through human experiences and interactions (Creswell & Poth, 2016). Constructivism enables researchers to understand how individuals interpret phenomena within their specific contexts. In the case of this study, it facilitates a nuanced exploration of how postgraduate supervisors perceive their readiness to integrate AI tools into the supervision process. The paradigm supports an interpretive lens through which the subjective meanings and experiences of supervisors are examined, making it particularly suitable for a study aimed at understanding human perspectives in a technologically evolving academic context (Lincoln et al., 2011).

The study adopts a qualitative research approach, which is appropriate for an in-depth exploration of complex human behaviours, beliefs, and experiences (Tisdell et al., 2025). A qualitative approach enables the collection of rich, descriptive data that captures the participants' views, attitudes, and experiences regarding the use of AI in postgraduate supervision. Within this approach, the study employs a generic qualitative research design. This design is valuable for uncovering patterns, generating insights, and developing a deeper understanding of how supervisors conceptualise and respond to AI integration (Babbie, 2020).

The population for this study consists of postgraduate supervisors in higher education institutions in South Africa. Using purposive sampling, 15 supervisors from diverse disciplines across multiple institutions (a mixture of public and private higher education institutions) were selected based on their direct involvement in supervising postgraduate students and their varying levels of familiarity with digital tools. This sampling method ensures that participants are information-rich and well-positioned to provide meaningful insights relevant to the research focus (Patton, 2015). The diversity in disciplinary backgrounds also adds depth to the findings by capturing cross-disciplinary variations in perceptions and readiness.

Data collection was conducted through semi-structured interviews, which allowed participants the flexibility to express their views while ensuring consistency in the key areas explored. Interviews were audio-recorded with participants' consent and subsequently transcribed for analysis. Thematic analysis was employed to interpret the data, following Braun and Clarke's (2006) six-phase framework: familiarisation with the data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the final report. This method supports the constructivist aim of interpreting meaning from participants' perspectives and facilitates the emergence of themes that reflect supervisors' readiness, perceived benefits and challenges, and contextual enablers or barriers to AI integration.

## 3. Presentation of Results and Discussion of Findings

The analysis of the interview data generated six interrelated themes, readiness across disciplines, AI integration into supervisory practices, perceptions and attitudes toward AI, institutional support and resources, technological literacy and familiarity, and workload and time constraints, which collectively illuminate the multifaceted factors shaping supervisors' readiness to integrate AI tools into postgraduate supervision.

### 3.1 Readiness across disciplines

The integration of AI tools into postgraduate supervision reveals varied levels of readiness among supervisors, shaped by their disciplinary orientations and the nature of research practices within their fields. Supervisors from more technically inclined disciplines expressed a higher level of readiness and familiarity with AI, while those from the humanities and social sciences adopted a more cautious or critical stance. Participant 4, an Engineering Supervisor, expressed a strong sense of readiness, noting that AI tools are already embedded in their research processes: "*AI tools have already become part of our research process, from data modelling to simulations, so for me, integrating them into supervision feels like a natural progression*." In contrast, Participant 5, a Law Supervisor, expressed uncertainty regarding the applicability of AI tools in their supervision practices: "*To be honest, I still struggle to see how AI fits into literary analysis or creative writing supervision. It feels a bit removed from our pedagogical approach*." Similarly, Participant 9, a Sociology Supervisor, acknowledged the potential efficiency AI could bring but voiced concerns about its implications for critical thinking in qualitative research: "*While I can appreciate the efficiency AI might bring to data analysis, I worry that it might dilute the critical thinking we aim to develop in qualitative research*." Participant 10, a Computer Science Supervisor, reported a high level of readiness, indicating that both they and their students are already engaging with AI tools in practical ways: "*My students are already using AI-based tools for code generation and testing, so as a supervisor, I have to stay ahead to guide them effectively*." From a more reflective standpoint, Participant 11, a Philosophy Supervisor, raised deeper epistemological concerns about the role of AI in academic supervision: "*There is a philosophical irony in relying on AI for thesis development. It challenges our very foundations about human reasoning and original thought*." Participant 13, an Interdisciplinary Supervisor, highlighted the conditional nature of their readiness, which depends on the AI tool in question and its alignment with educational objectives: "*Our field sits between the technical and the humanistic, so my readiness depends on the specific AI tool and how it aligns with the educational goals of my students*."

The data underscores that disciplinary orientations significantly influence the level of familiarity, openness, and critical engagement with AI technologies in supervisory practices. Supervisors in technical and STEM-related disciplines, such as Engineering and Computer Science, demonstrate a high degree of readiness and fluency with AI tools. This observation aligns with existing literature suggesting that disciplines engaged in computational, quantitative, and data-driven research are more predisposed to adopt emerging technologies, including AI, due to their embeddedness in established research workflows (Selten & Klievink et al., 2024). Participant 4's observation reflects this alignment, where AI's role in modelling, simulations, and automation seamlessly integrates with disciplinary norms. Similarly, Participant 10 from Computer Science highlighted that AI is not merely a supervisory tool but also a learning companion for students, particularly in programming, code generation, and software testing. This supports the argument put forth by Kamalov et al. (2023) that AI has already penetrated technical education landscapes, becoming indispensable in both instruction and assessment.

In contrast, supervisors from the humanities and social sciences articulated more caution and critical reflection, revealing a tension between AI integration and disciplinary epistemologies. Participant 5's concern regarding the applicability of AI in literary analysis echoes anxieties within the humanities about the potential reduction of interpretative depth and the threat to creativity and originality (Yadav, 2024). These concerns are compounded by Participant 9's comment on critical thinking, highlighting a fear that AI-driven efficiency may inadvertently erode cognitive rigour, a cornerstone of qualitative inquiry. Participant 11's philosophical interrogation of AI's epistemological implications raises an important point about the role of AI in reshaping academic values. This response draws attention to deeper ontological questions surrounding authorship, intellectual agency, and the authenticity of student work—issues also echoed in the work of Memarian and Doleck (2023), who argue that AI in higher education may inadvertently commodify learning processes.

Interestingly, the interdisciplinary perspective offered by Participant 13 reveals that readiness is not always binary but rather conditional, shaped by how AI tools align with educational intentions. This nuanced stance supports the findings of Uren and Edwards (2023), who advocate for a cross-disciplinary approach in integrating AI, emphasising that adaptability and critical alignment are necessary for meaningful adoption. Collectively, the data suggest that disciplinary cultures, values, and methodologies deeply inform supervisors' readiness to integrate AI. As Fu and Weng (2024) contend, educational technology cannot be divorced from the contexts in which it is applied; disciplinary traditions must be central to conversations about AI adoption in postgraduate supervision.

### 3.2 AI Integration into supervisory practices

Supervisors' readiness to integrate AI into postgraduate supervision is reflected in their current usage, ongoing exploration, and imaginative projections of AI's potential. The data captures both the practical adoption of AI tools in everyday supervision and the supervisors' openness to future integration. Their experiences demonstrate a spectrum, from current application to future possibilities, which indicates a growing awareness and readiness to embed AI meaningfully into academic support. Participant 1 shared how plagiarism detection and writing enhancement tools have already become routine in their supervision practices: "*I already use Turnitin and Grammarly with my students; it has become part of my normal workflow. It saves time and ensures that we catch issues early, especially with referencing and originality.*" Expanding on current uses, Participant 7, a supervisor from the Social Sciences, described their recent engagement with more advanced AI tools to support students in developing their literature reviews and refining research questions: "*Recently, I have started experimenting with AI tools like Elicit and Scite to help students frame their literature reviews and refine their research questions. These tools are game changers, especially for first-time researchers.*" Participant 6 highlighted how AI contributes to methodological clarity for students, especially those unfamiliar with research designs: "*I have used AI-based platforms to help students explore appropriate methods for their research. It does not replace my input, but it is a good starting point, especially for those unfamiliar with qualitative or mixed methods designs.*" Some supervisors also pointed to the potential of AI to enhance communication and feedback mechanisms. Participant 8 envisioned a department-level AI assistant to handle routine student queries: "*Imagine if we had an AI chatbot for our department that could answer basic questions about formatting or deadlines; students wouldn't always have to wait for an email reply. That kind of support would really enhance the supervision process.*" Participant 15 echoed this optimism, focusing on the role of AI in streamlining feedback: "*AI can provide instant feedback on students' writing, especially in early drafts. This could free me up to focus on the deeper intellectual engagement during one-on-one meetings.*" Participant 12 acknowledged their exploratory phase, viewing AI as a promising enhancement rather than a replacement for their academic role: "*While I am not fully integrating AI in all aspects of supervision yet, I am actively exploring its potential. I see it as a complementary tool, not a threat; it is about adapting to what benefits the student most.*"

The data shows that AI integration is already underway, with many supervisors incorporating it into their teaching practices. Their growing comfort and proactive engagement highlight a shift from mere awareness to imaginative, experience-based application. Participant 1's routine use of Turnitin and Grammarly illustrates the mainstreaming of basic AI tools in academic supervision. While these tools are not generative AI, they serve foundational roles in promoting academic integrity and writing quality. Their widespread adoption reflects AI's integration into core educational workflows, consistent with Malik et al. (2023), who observed that text-based AI systems are often the initial entry points for AI adoption in higher education settings. They offer efficiency, early feedback, and reduced administrative burden, helping supervisors manage quality assurance with greater ease.

The use of more sophisticated tools, such as Elicit and Scite, as described by Participant 7, points to an emerging shift towards cognitive augmentation, where AI supports higher-order academic tasks like literature synthesis and question refinement. These tools leverage natural language processing and evidence-based AI to help students navigate vast academic databases and identify relevant scholarly arguments. According to Bayly-Castaneda et al. (2024), AI that enhances cognitive engagement—not just procedural tasks—marks the next frontier in educational technology, allowing for more personalised and efficient learning journeys. Participant 6's use of AI for methodological guidance emphasises AI's potential as a pedagogical scaffold. Supervisors are no longer solely bearers of content but facilitators of AI-mediated exploration. This resonates with Adewale et al. (2024), who argue that AI can democratise access to complex academic knowledge, helping students better understand methodological options and research design choices.

While human supervision remains irreplaceable for critical discussion and ethical reasoning, AI can serve as a springboard for inquiry, especially for novice researchers. Forward-looking perspectives, such as Participant 8's idea of an AI chatbot for departmental queries, reflect an imaginative but practical vision of AI's administrative and communicative utility. As Dakakni and Safa (2023) note, AI-powered virtual assistants are becoming viable solutions in educational settings to handle repetitive, non-cognitive tasks. Their integration could significantly enhance responsiveness and student satisfaction, freeing supervisors to focus on mentorship and academic rigour. Similarly, Participant 15's emphasis on automated writing feedback aligns with literature that champions adaptive feedback systems. AI-enabled feedback, especially in early drafts, can promote self-directed learning and reduce delays in formative assessment cycles. However, as Zhai et al. (2024) caution, such systems must be carefully calibrated to avoid over-reliance on machine-generated evaluation, which can depersonalise the feedback process if not complemented with human insight.

### 3.3 Perceptions and attitudes toward AI

Supervisors' perceptions and attitudes toward AI tools significantly influence their readiness to adopt such technologies in postgraduate supervision. While several supervisors viewed AI as a complementary resource that enhances efficiency and support, others expressed caution, highlighting concerns about academic integrity, ethical risks, and institutional preparedness. Participant 1, who supervises multiple students simultaneously, regarded AI as a helpful assistant in managing feedback more efficiently: "*I see AI as a valuable assistant; it helps me speed up the initial feedback process, especially when I have several students submitting drafts at the same time.*" Similarly, Participant 5 reflected positively on the role of AI in refining research focus and streamlining supervision: "*Using AI to help identify relevant literature or refine students' research questions has made supervision more focused and efficient. It is like having a co-pilot.*" Participant 12 appreciated AI's potential in supporting individualised student needs, especially in academic writing: "*With AI, I can tailor support for each student's writing needs. It does not replace my input, but it certainly helps in scaffolding their progress.*" However, not all supervisors shared the same level of enthusiasm. Participant 3 expressed concern about students becoming overly dependent on AI, potentially undermining the development of critical thinking skills: "*My concern is that students might start leaning too heavily on AI tools and bypass the critical thinking process entirely. That would be detrimental to their learning.*" Participant 10 voiced hesitation around the lack of clear institutional policies, particularly regarding ethical use and plagiarism: "*I am hesitant to promote AI tools until there is a clear policy on* *academic integrity and plagiarism. The line between assistance and misconduct is blurry.*" In addition, Participant 4 raised issues around data privacy and institutional readiness, noting the absence of structured guidance on responsible AI use: "*We have not had any proper training on how to use these tools responsibly. I worry about data privacy and what happens to students' work once it goes into the system.*"

Participants 1 and 5 characterised AI as a supportive co-facilitator in managing extensive supervision loads, refining research questions, and streamlining the feedback process. These positive attitudes are consistent with findings by Tan et al. (2024), who argue that AI can mitigate administrative and cognitive burdens on educators by automating repetitive tasks and enhancing decision-making processes. When AI is conceptualised as a co-pilot or assistant, as articulated by Participant 5, supervisors are more inclined to perceive it as an augmentative rather than a substitutive force (Spring et al., 2022). Participant 12 further emphasised AI's capability to personalise feedback and scaffold student writing, reflecting a belief in AI as a tool for personalised learning, a notion widely supported in the literature. According to Tan et al. (2024), AI technologies can facilitate adaptive learning experiences that respond to the individual needs of students, thereby meaningfully complementing human supervision.

However, this optimism is tempered by concerns expressed by others. For instance, Participant 3 expressed apprehension regarding over-reliance on AI and its potential to undermine critical thinking. This concern aligns with the warning from Delello et al. (2025), who caution that while AI can streamline learning, it risks fostering superficial engagement with content if not critically mediated by educators. Participant 10's reservations about institutional unpreparedness and ethical ambiguity highlight a significant challenge: the absence of clear guidelines governing the use of AI in higher education. This uncertainty may impede adoption, as supervisors remain uncertain about how to address issues such as plagiarism, authorship, and the ambiguous boundaries between support and misconduct (Memarian & Doleck, 2023). Participant 10's advocacy for institutional policy resonates with Nguyen et al. (2023), who argue that successful integration of AI in academia must be supported by robust ethical frameworks and institutional support systems. Moreover, Participant 4's concerns regarding data privacy and responsible use draw attention to an often-overlooked aspect of AI deployment: the governance of student data. As Mienye and Swart (2025) contend, the collection and processing of student-generated data raise serious concerns regarding consent, transparency, and surveillance in educational settings. Without structured guidance or training, supervisors may hesitate to engage with AI tools that could jeopardise academic freedom or students' intellectual property.

### 3.4 Institutional support and resources

Participants indicate that institutional support plays a pivotal role in shaping supervisors' readiness to integrate AI tools into postgraduate supervision. The availability of resources, structured training, and clear frameworks significantly influences how confident and prepared supervisors feel. Participant 11 emphasised that access to institutional AI tools instils confidence in both supervisors and students: "*Our university has made AI tools like Grammarly and Turnitin available, and that kind of institutional backing makes it easier for me to introduce these tools to my students with confidence.*" Participant 9 highlighted the importance of consistent encouragement and hands-on learning opportunities provided by the institution: "*We are constantly encouraged to explore digital innovations, and the institution provides regular workshops on tools like ChatGPT. That support really boosts my confidence in using AI effectively.*" From a structural and departmental level, Participant 6 appreciated the integration of digital literacy within the supervision framework, describing it as a key factor in their preparedness: "*I feel more prepared because our department integrates digital literacy and AI awareness into our postgraduate supervision strategy. That kind of systemic support matters.*" On the other hand, Participant 14 expressed concern about the absence of formal guidance from the institution, stressing the burden placed on individual supervisors: "*If the institution does not create a clear framework for AI use in supervision, it puts too much responsibility on individual supervisors to figure it out on their own.*"

The data highlights that access to technological tools alone is insufficient; what truly enhances readiness is a comprehensive, systemic framework that includes training, policy guidance, and a culture of innovation. Participant 11 credited their confidence in adopting AI tools to the availability of institutional resources such as Grammarly and Turnitin. Although these tools are not generative AI in the strictest sense, they illustrate how institutional endorsement lends credibility to the use of AI in supervision. This finding aligns with those of Zawacki-Richter et al. (2019), who emphasise that when AI tools are institutionally sanctioned, supervisors feel more secure regarding their legitimacy and pedagogical appropriateness. The endorsement of tools by universities indicates that such technologies have passed ethical and academic scrutiny, thereby reducing the risk for individual supervisors. Participant 9's reference to hands-on workshops and encouragement to explore innovations underscores the importance of ongoing professional development.

As Chan and Hu (2023) argue, the effective integration of AI in higher education depends not only on the availability of tools but also on the surrounding training ecosystem. Institutional readiness, therefore, must encompass consistent investment in capacity-building initiatives that assist academic staff in developing competence and confidence. Participant 6's experience with digital literacy embedded within departmental frameworks highlights the value of aligning AI integration with broader curricular and strategic priorities. Tan et al. (2016) stress that the systemic incorporation of digital and AI literacy in academic programmes ensures that both supervisors and students operate within a coherent, future-oriented educational vision. This proactive institutional stance fosters a shared culture of AI engagement, rather than relegating it to isolated initiatives or tech-savvy individuals. However, Participant 14's concern regarding the absence of clear institutional guidance underscores a critical gap. Without formal policies or frameworks, the burden of AI integration disproportionately falls on individual supervisors, leading to inconsistencies in usage and potentially conflicting interpretations of academic integrity. This concern is echoed in Akgun and Greenhow (2022), who argue that institutions must take responsibility for the ethical, pedagogical, and technical implications of AI, rather than simply delegating them to frontline educators. Furthermore, Makore (2024) highlights the risk of creating technological patchworks when AI adoption is left to individual experimentation. In such contexts, some supervisors may thrive while others remain disengaged, resulting in uneven learning experiences for students.

### 3.5 Technological literacy and familiarity

Technological literacy and familiarity emerged as significant factors influencing supervisors' readiness to integrate AI tools into postgraduate supervision. Supervisors with prior experience using digital tools or those from technology-driven disciplines expressed greater confidence and enthusiasm regarding AI integration. Conversely, those with limited digital exposure or competing professional responsibilities reported feelings of apprehension and resistance. Participant 13, an interdisciplinary supervisor, expressed comfort with AI integration due to their existing familiarity with digital tools: "*Because I already use Turnitin and some basic data analysis software in my supervision process, I find the idea of integrating AI tools quite natural. I see it as an extension or advancement of what I am already doing*." Similarly, Participant 15, from a Computer Science background, highlighted how their discipline fosters technological confidence: "*I come from a Computer Science background, so I am quite familiar with digital innovations and operations. AI does not intimidate me; I see it as an opportunity to improve supervision efficiency*." Participant 4, a STEM supervisor, attributed readiness to generational familiarity and discipline-specific exposure to technological change: "*The younger generation of supervisors, especially in STEM fields, are generally more open to using technology. We have grown alongside it. For us, incorporating AI feels more like an evolution than a disruption*." In contrast, supervisors from non-technical disciplines expressed discomfort and concern. Participant 8, from the humanities, described feeling overwhelmed by digital tools: "*To be honest, I am from a humanities background, and I have never been comfortable with digital tools beyond email and Word. The idea of using AI in supervision is quite overwhelming. I am not familiar with it*." For Participant 14, the challenge was not a lack of awareness but a lack of time and capacity to learn new systems amid existing responsibilities: "*It is not that I do not see the value of AI, but I just do not have the time to learn another tool right now. Between teaching, admin, and supervision, it feels like an extra burden*." Finally, Participant 2 raised philosophical and relational concerns, emphasising a fear of losing meaningful human interaction in the supervisory process: "*There is a fear of the unknown. Some of us worry that relying on AI might compromise the depth of human interaction that is essential in postgraduate supervision*."

Participant 13's reflection that AI is a "natural extension" of tools such as Turnitin and data analysis software illustrates how familiarity breeds confidence. This aligns with Zawacki-Richter et al. (2019), who argue that prior engagement with educational technologies fosters a positive disposition towards adopting new tools. For such supervisors, AI is not a radical departure but an enhancement of current practices, which Collins et al. (2021) describe as technology-enhanced learning ecosystems. Similarly, Participants 15 and 4, both from Computer Science and STEM backgrounds, emphasised how disciplinary training and generational exposure to digital environments create a culture of technological adaptability. Their readiness reflects the argument made by Malik et al. (2023) that digital natives and those in computational disciplines are more likely to embrace emerging technologies due to the alignment between their workflows and AI's functionalities, such as automation, modelling, and data analysis.

However, supervisors in the humanities and non-technical fields presented a contrasting view. Participant 8's comment about struggling beyond basic tools such as email or Word reveals a technological literacy divide, a barrier also noted by Rioseco-Pais et al. (2024), who caution against assuming uniform digital competence across academia. This divide not only affects AI adoption but also risks exacerbating inequalities in pedagogical innovation. Participant 14's perspective adds a pragmatic dimension to this barrier, which is not rooted in resistance but in time and capacity constraints. The competing demands of teaching, administration, and supervision often make professional upskilling seem burdensome. This is consistent with Kaputa et al. (2022), who emphasise that digital transformation in higher education is not just a technological issue but also a matter of institutional support and workload management. Participant 2's philosophical concern about AI undermining human interaction touches on a relational and pedagogical dilemma. The fear that AI might depersonalise supervision or reduce the richness of mentorship aligns with critiques by Zhai et al. (2024), who warn against the depersonalisation of education in the face of automation. In supervision, a process rooted in trust, dialogue, and intellectual growth, this concern cannot be overlooked.

### 3.6 Workload and time constraints

The data revealed that one of the key factors influencing supervisors' readiness to integrate AI tools into postgraduate supervision is the constraint of time and their existing workload. Many supervisors recognise the potential benefits of AI in alleviating certain aspects of their responsibilities; however, the significant demands of teaching, research, and administrative duties permit little opportunity to explore or adopt new technologies. Participant 11 emphasised how their multifaceted academic responsibilities significantly hinder their ability to engage with new technologies: "*I have back-to-back responsibilities such as teaching, administration, and supervision, which affect my readiness for the adoption and integration of AI tools into my students' supervision. Honestly, I do not have the time to explore or learn how AI tools work, even if they might help in the long run.*" Similarly, Participant 1 noted that while the promise of AI for automating feedback is appealing, the time investment required to learn a new system is a major barrier: "*The idea of AI sounds promising, especially for feedback automation, but learning a new system takes time I just do not have right now.*" Participant 12 shared comparable sentiments, recognising the potential of AI tools but expressing concern about the learning curve amidst their current commitments: "*I think AI could reduce my marking load or help with literature suggestions, but the initial learning curve feels overwhelming considering my current workload.*" Participant 7 echoed the general sense of overload experienced by supervisors, suggesting that AI adoption currently feels more like an additional burden than a solution: "*We are already overloaded. Adding AI training on top of everything else feels like another task rather than a solution.*" Participant 14 pointed out that institutional support must go beyond training and include schedule adjustments if AI is to be realistically integrated into supervisory practices: "*If institutions really want us to use AI, they need to make space in our schedules. As it stands, we are too stretched to take on anything new.*"

Participant 11's reflection encapsulates a prevalent reality in higher education: supervisors are overextended across teaching, research, and administration. The tension between these responsibilities mirrors findings by Rasool et al. (2022), who argue that academic workloads are increasingly fragmented and intense, leaving little time for professional upskilling or technological experimentation. For many supervisors, the adoption of artificial intelligence (AI) is not a matter of resistance to innovation but rather a consequence of limited temporal bandwidth. The comments from Participant 1 and Participant 12 highlight the paradox of technological adoption. AI is perceived as a tool that could ultimately reduce workload (e.g., automating feedback, generating literature suggestions), yet the initial investment of time required to learn and adapt creates a barrier to adoption.

This challenge is echoed in Weller's (2020) critique of innovation in academia, wherein technological tools are often introduced without sufficient time or structural support for their integration. Participant 7's observation that AI feels like "another task" rather than a solution reinforces this notion. When innovations are imposed without alleviating existing obligations, they risk being perceived as additional burdens, regardless of their long-term value. Furthermore, the concern raised by Participant 14 regarding institutional support signals a systemic issue. Institutional ambitions for AI-enhanced supervision must be matched with structural changes, such as workload reallocation, dedicated training time, or even incentivised professional development. This aligns with Zawacki-Richter et al. (2019), who emphasise that AI readiness in higher education must be supported through strategic leadership, policy alignment, and capacity-building initiatives. Without such efforts, individual readiness will remain limited, and AI tools will likely remain underutilised despite their potential.

## 4. Conclusions

Drawing explicitly from the empirical findings of this study, the following conclusions are proposed. Supervisors' readiness to integrate artificial intelligence (AI) is uneven and dependent on discipline. The findings indicate that this readiness is significantly influenced by disciplinary orientation. Supervisors within STEM and technically oriented disciplines exhibit higher levels of readiness, primarily because AI tools are already embedded in their research workflows and epistemic traditions. In contrast, supervisors in the humanities and social sciences display a more cautious engagement, rooted in concerns regarding interpretive depth, originality, critical thinking, and epistemological integrity. This confirms that readiness is not a generic competency but is mediated by disciplinary norms and knowledge practices.

AI integration is already occurring, albeit at differentiated levels of complexity. Evidence from the findings illustrates that supervisors are engaging with AI tools along a continuum. At the foundational level, tools such as Turnitin and Grammarly are widely normalised and incorporated into routine supervision practices. More advanced engagement, including the utilisation of tools such as Elicit and Scite for literature development and research framing, is emerging but remains inconsistent. This suggests that readiness manifests not as full adoption or rejection, but as incremental and layered engagement shaped by familiarity and perceived usefulness. Supervisors largely conceptualise AI as an augmentative rather than substitutive force. Across disciplines, supervisors consistently position AI as a support mechanism rather than a replacement for human supervision. The findings indicate that AI is valued for its capacity to reduce administrative burdens, support early-stage writing and literature work, and enhance efficiency, while core supervisory functions such as intellectual mentoring, ethical judgement, and epistemic guidance remain firmly centred on human input. This framing is central to supervisors' willingness to engage with AI.

Perceptions of artificial intelligence (AI) are influenced by ethical uncertainties and concerns regarding excessive reliance on technology. While several supervisors conveyed optimism about the pedagogical potential of AI, the findings also reveal enduring anxieties surrounding academic integrity, student dependency, data privacy, and the ambiguities between assistance and misconduct. These concerns restrict readiness, particularly in the absence of clear institutional guidance. Consequently, supervisors' attitudes towards AI are ambivalent, reflecting both opportunities and risks rather than outright resistance. Institutional support emerges as a critical enabler of readiness. The findings clearly indicate that supervisors operating within institutions that offer access to approved AI tools, structured training, and explicit policy frameworks report increased confidence and willingness to integrate AI into their supervisory practices. In contrast, in the absence of institutional guidance, supervisors encounter uncertainty and bear the ethical and practical burdens individually. Readiness is thus not solely an individual attribute but is profoundly contingent upon institutional culture, leadership, and infrastructure.

Technological literacy and workload pressures significantly influence readiness. Supervisors with prior digital exposure exhibit greater confidence and openness towards AI integration, while those with limited technological familiarity express anxiety and resistance. However, even technologically adept supervisors acknowledge that workload and time constraints serve as major barriers. The findings suggest that readiness is weakened when the adoption of AI is introduced without corresponding adjustments to workload or provisions for protected time for learning. This study concludes that supervisors' readiness to integrate AI into postgraduate supervision is contingent rather than static. It is shaped by the intersection of disciplinary epistemologies, technological familiarity, ethical positioning, institutional support, and workload realities. Effective integration of AI in postgraduate supervision thus necessitates systemic, context-sensitive approaches rather than one-size-fits-all solutions.

### 4.1 Recommendations

The study recommends that higher education institutions implement ongoing, discipline-sensitive professional development programmes that equip supervisors with AI competencies in ways that respect and enhance their existing research traditions. Such initiatives carry significant social implications, as they can reduce disparities between supervisors who are digitally confident and those who feel marginalised by rapid technological change. Tailored training, coordinated by academic development units in collaboration with universities, can help ensure that supervisors across different age groups, backgrounds, and disciplines benefit equitably from AI-enhanced supervision practices.

Embedding ethical, inclusive, and socially responsible AI literacy within these programmes can further support supervisors in navigating issues of authorship, academic integrity, and student dependency. These efforts have practical implications for improving the quality of postgraduate supervision by strengthening supervisors' ability to provide consistent, timely, and pedagogically sound feedback. Institutions should also foster interdisciplinary dialogue through communities of practice, workshops, and structured forums where supervisors can collectively reflect on challenges, share experiences, and exchange strategies. Such platforms not only support the practical development of AI skills but also promote a reflective academic culture that enhances collegiality, reduces isolation, and bridges disciplinary divides in AI adoption.

At the institutional level, university leadership, together with research, legal, and ethics offices, should develop comprehensive policies governing AI use in postgraduate supervision. Clear guidelines on academic integrity, authorship, privacy, and responsible usage will ensure that both supervisors and students operate within safe and transparent boundaries. The social implication of robust policy frameworks lies in building trust within the academic community, ensuring that AI adoption does not inadvertently compromise equity, fairness, or intellectual integrity. Practically, these policies will support consistent decision-making across departments and reduce the burden on individual supervisors to interpret ethical grey areas independently.

Institutions should also invest in academically vetted AI tools and integrate them into existing digital ecosystems. ICT departments, in collaboration with relevant academic units, should ensure that supervisors have access to reliable tools and ongoing technical support. This investment has clear practical benefits, such as improving efficiency in feedback provision, supporting data management, and enhancing the overall supervision experience for both supervisors and students. Additionally, institutions should implement monitoring and evaluation mechanisms to assess the impact of AI on supervision practices. Feedback gathered from supervisors and students can guide iterative improvements to tools, training, and policy frameworks, ensuring that AI integration remains sustainable, ethically sound, and pedagogically meaningful.

Based on the study, the following recommendations for further inquiry are proposed: Future research could investigate how discipline-specific pedagogical values and epistemologies influence supervisors' adoption and critical engagement with AI tools in postgraduate supervision. Additional studies could examine the effectiveness of institutional frameworks, such as training programmes and policy support, in enhancing supervisors' digital literacy and ethical preparedness for AI integration.

## 5. Declarations

**Funding:** This research did not receive any external funding.
**Conflicts of Interest:** The author declares no conflict of interest.
**Use of Artificial Intelligence:** The current work was created without the assistance of artificial intelligence technologies, as confirmed by the author.

## References

Adewale, M. D., Azeta, A., Abayomi-Alli, A., & Sambo-Magaji, A. (2024). Impact of artificial intelligence adoption on students' academic performance in open and distance learning: A systematic literature review. *Heliyon, 10*, 1–19. https://doi.org/10.1016/j.heliyon.2024.e40025

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics, 2*(3), 431–440. https://doi.org/10.1007/s43681-021-00096-7

Babbie, E. R. (2020). *The practice of social research* (15th ed.). Cengage Learning.

Balalle, H., & Pannilage, S. (2025). Reassessing academic integrity in the age of AI: A systematic literature review on AI and academic integrity. *Social Sciences & Humanities Open, 11*, 1–12. https://doi.org/10.1016/j.ssaho.2025.101299

Bayly-Castaneda, K., Ramirez-Montoya, M. S., & Morita-Alexander, A. (2024). Crafting personalised learning paths with AI for lifelong learning: A systematic literature review. *Frontiers in Education, 9*, 1–12. https://doi.org/10.3389/feduc.2024.1424386

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*(1), 1–18. https://doi.org/10.1186/s41239-023-00411-8

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management, 60*, 1–17. https://doi.org/10.1016/j.ijinfomgt.2021.102383

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage.

Dakakni, D., & Safa, N. (2023). Artificial intelligence in the L2 classroom: Implications and challenges on ethics and equity in higher education: A 21st-century Pandora's box. *Computers and Education: Artificial Intelligence, 5*, 1–10. https://doi.org/10.1016/j.caeai.2023.100179

Delello, J. A., Sung, W., Mokhtari, K., Hebert, J., Bronson, A., & De Giuseppe, T. (2025). AI in the classroom: Insights from educators on usage, challenges, and mental health. *Education Sciences, 15*(2), 1–27. https://doi.org/10.3390/educsci15020113

Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centred AI practices. *Computers and Education: Artificial Intelligence, 7*, 1–20. https://doi.org/10.1016/j.caeai.2024.100306

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New era of artificial intelligence in education: Towards a sustainable multifaceted revolution. *Sustainability, 15*(16), 1–27. https://doi.org/10.3390/su151612451

Kaputa, V., Loučanová, E., & Tejerina-Gaite, F. A. (2022). Digital transformation in higher education institutions as a driver of socially oriented innovations. In *Social innovation in higher education* (pp. 81–85). Springer. https://doi.org/10.1007/978-3-030-84044-0_4

Kassa, B. Y., & Worku, E. K. (2025). The impact of artificial intelligence on organizational performance: The mediating role of employee productivity. *Journal of Open Innovation: Technology, Market, and Complexity, 11*, 1–17. https://doi.org/10.1016/j.joitmc.2025.100474

Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F., ... & Palmer, E. (2024). The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers and Education: Artificial Intelligence, 6*, 1–10. https://doi.org/10.1016/j.caeai.2024.100221

Lincoln, Y. S., Lynham, S. A., & Guba, E. G. (2011). Paradigmatic controversies, contradictions, and emerging confluences, revisited. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (4th ed., pp. 97–128). Sage.

Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., & Darwis, A. (2023). Exploring artificial intelligence in academic essay: Higher education students' perspective. *International Journal of Educational Research Open, 5*, 1–11. https://doi.org/10.1016/j.ijedro.2023.100296

Makore, S. T. (2024). Regulating artificial intelligence to advance financial inclusion in South Africa. *Potchefstroom Electronic Law Journal (PELJ), 27*(1), 1–35. https://doi.org/10.17159/1727-3781/2024/v27i0a17488

Memarian, B., & Doleck, T. (2023). Fairness, accountability, transparency, and ethics (FATE) in artificial intelligence (AI) and higher education: A systematic review. *Computers and Education: Artificial Intelligence, 5*, 1–12. https://doi.org/10.1016/j.caeai.2023.100152

Mienye, I. D., & Swart, T. G. (2025). ChatGPT in education: A review of ethical challenges and approaches to enhancing transparency and privacy. *Procedia Computer Science, 254*, 181–190. https://doi.org/10.1016/j.procs.2025.01.022

Mulenga, R., & Shilongo, H. (2024). Academic integrity in higher education: Understanding and addressing plagiarism. *Acta Pedagogia Asiana, 3*(1), 30–43. https://doi.org/10.53623/apga.v3i1.337

Nguyen, A., Kremantzis, M., Essien, A., Petrounias, I., & Hosseini, S. (2024). Enhancing student engagement through artificial intelligence (AI): Understanding the basics, opportunities, and challenges. *Journal of University Teaching and Learning Practice*, *21*(6), 1-13.

Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies, 28*(4), 4221–4241. https://doi.org/10.1007/s10639-022-11316-w

Nouri, J., Zhang, L., Mannila, L., & Norén, E. (2020). Development of computational thinking, digital competence and 21st century skills when learning programming in K–9. *Education Inquiry, 11*(1), 1–17. https://doi.org/10.1080/20004508.2019.1627844

Omodan, B. I. (2025). Redefining the role of supervisors in the era of artificial intelligence: Implications for hybrid postgraduate research governance. *Cogent Education, 12*(1), 1–16. https://doi.org/10.1080/2331186X.2025.2536534

Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Sage.

Rasool, T., Warraich, N. F., & Sajid, M. (2022). Examining the impact of technology overload at the workplace: A systematic review. *Sage Open, 12*(3), 1–18. https://doi.org/10.1177/21582440221114320

Rasul, T., Nair, S., Kalendra, D., Balaji, M. S., de Oliveira Santini, F., Ladeira, W. J., ... & Hossain, M. U. (2024). Enhancing academic integrity among students in the GenAI era: A holistic framework. *The International Journal of Management Education, 22*(3), 1–14. https://doi.org/10.1016/j.ijme.2024.101041

Rioseco-Pais, M., Silva-Quiroz, J., & Vargas-Vitoria, R. (2024). Digital competences and years of access to technologies among Chilean university students: An analysis based on the DIGCOMP framework. *Sustainability, 16*(22), 1–14. https://doi.org/10.3390/su16229876

Sehmi, A., Sarfraz, I., & Hussain, M. (2025). Generative artificial intelligence's integration for data analysis in conducting academic research: Understanding the perspective of research supervisors. *Journal of Advanced Academics, 36*(4), 788–815. https://doi.org/10.1177/1932202X251365312

Selten, F., & Klievink, B. (2024). Organising public sector AI adoption: Navigating between separation and integration. *Government Information Quarterly, 41*(1), 1–14. https://doi.org/10.1016/j.giq.2023.101885

Spring, M., Faulconbridge, J., & Sarwar, A. (2022). How information technology automates and augments processes: Insights from artificial-intelligence-based systems in professional service operations. *Journal of Operations Management, 68*(6–7), 592–618. https://doi.org/10.1002/joom.1215

Tan, X., Cheng, G., & Ling, M. H. (2024). Artificial intelligence in teaching and teacher professional development: A systematic review. *Computers and Education: Artificial Intelligence, 8*, 1–19. https://doi.org/10.1016/j.caeai.2024.100355

Tisdell, E. J., Merriam, S. B., & Stuckey-Peyrot, H. L. (2025). *Qualitative research: A guide to design and implementation.* John Wiley & Sons.

Uren, V., & Edwards, J. S. (2023). Technology readiness and the organisational journey towards AI adoption: An empirical study. *International Journal of Information Management, 68*, 1–12. https://doi.org/10.1016/j.ijinfomgt.2022.102588

Weller, M. (2020). *25 years of ed tech.* Athabasca University Press.

Yadav, D. (2024). The role of artificial intelligence in literary analysis: A computational approach to understand literary styles. *International Journal of Emerging Knowledge Studies, 3*(9), 558–565. https://doi.org/10.70333/ijeks-03-09-006

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*(1), 1–37. https://doi.org/10.1186/s40561-024-00316-7

# Supervision Skills of Supervisors in AI-enhanced Environments: Perspectives on Postgraduate Supervision

**Thembi B. Nkosi**[1]

AFFILIATIONS

[1]Provincial Department of Education: Quality Assurance (Quality Management and Support), Mpumalanga Department of Education, South Africa.

CORRESPONDENCE

Email: nkosi.thembi1@gmail.com

REFERENCE

Nkosi, T. B. (2026). Supervision skills of supervisors in AI-enhanced environments: Perspectives on postgraduate supervision. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 93-110). ERRCD Forum. https://doi.org/10.38140/obp4-2026-06



**Abstract:** The use of AI by postgraduate students quickly changes supervisory relationships and requires new supervisory skills. This study examines the fundamental supervisory abilities needed to manage postgraduate students who integrate AI tools into their research work. It employs a qualitative research method based on exploratory phenomenology within an interpretive research paradigm to investigate supervisors' subjective experiences and perspectives in AI-integrated supervision environments. Ten purposively selected supervisors with experience in AI-enhanced settings provided data through semi-structured interviews. An analysis of the interview transcripts using thematic methods revealed consistent patterns and themes regarding supervisory competencies. Supervisors need to cultivate critical evaluation skills to identify students' overdependence on AI systems and learn how to detect AI-generated material that lacks originality by interpreting underlying meanings. Students require guidance from supervisors in learning essential research techniques, such as literature searching and correct source attribution, to uphold academic integrity. The study emphasises the importance of supervisors mandating students to record their research steps and participate in evaluative discussions to test their understanding and ethical use of AI. Supervisory responsibilities must incorporate AI tools while simultaneously promoting independent critical thought and ethical principles. The study proposes specialised training programmes for supervisors to enhance their AI literacy and evaluation skills while also creating clear ethical guidelines for AI use in postgraduate research. Future research should investigate how AI integration affects supervisory relationships over time and develop scalable supervisor training frameworks suitable for various academic fields and institutional settings.

***Keywords:*** Artificial intelligence, digital literacy, higher education, human-AI collaboration, postgraduate supervision, supervisory skills.

## 1. Introduction

AI integration in professional settings has significantly changed supervisory functions, especially within postgraduate supervision frameworks. The development of advanced AI systems that handle complex tasks has altered supervisors' traditional responsibilities. This chapter investigates the necessary supervisory skills for supervisors working in AI-powered environments. Effective supervision throughout history has depended on supervisors who possess deep knowledge and expertise in their tasks to guide and mentor students appropriately (Corey et al., 2020; Davys & Beddoe, 2020; Lee, 2019). Supervisory roles have transformed due to advancements in AI tools, resulting in new supervisors who often lack firsthand experience in task performance. The emergence of AI-driven operations raises essential inquiries about the preparation needed for supervisors who lack direct manual task execution experience (Bainey, 2024; Bhadoriya, 2024; Le, 2024). The current body of research identifies multiple challenges faced when preparing supervisors to work in AI-enhanced settings. A learning gap emerges when supervisors have no prior experience with tasks before AI integration (Rajabi, 2023). The existence of this gap limits supervisors' capacity to solve problems and drive innovation within AI systems. Organisations should fund extensive training initiatives that develop supervisory skills for AI technology management while promoting a continuous learning atmosphere (George & Wooden, 2023; Jarrahi et al., 2023; Morandini et al., 2023). The Four Pillars of Oversight and Four Steps of Supervisory Flow frameworks illustrate how supervisors can integrate AI tools effectively while preserving necessary human oversight (Kyriakou & Otterbacher, 2023; Mack, 2023). Supervisory frameworks focus on oversight, improvement, empowerment, and harmonisation to ensure the ethical and productive use of AI tools. Postgraduate supervision requires a clear understanding and development of supervisory skills due to its growing intersection with AI technologies. Through targeted professional development and addressing educational deficiencies, educational institutions can equip supervisors to succeed in the changing academic landscape and enhance the quality of postgraduate education through AI integration.

AI adoption in professional work environments has altered supervisory practices across various settings, including postgraduate supervision. The development of AI systems for the automated performance of complex tasks has shifted the original functions of supervisors. This chapter examines the specific supervisory abilities that supervisors need to possess when working in AI-enhanced environments. Effective supervisory practices of the past required supervisors to have task-specific knowledge and competency to guide students through similar tasks (Corey et al., 2020; Davys & Beddoe, 2020).

Due to advancements in AI toolkits, supervisors have evolved, and new supervisors are often hired without prior experience in the tasks performed by the organisation. AI-enabled operations raise urgent questions about how to prepare supervisors who lack a background in the manual execution of tasks (Bainey, 2024). The existing literature on the topic has identified several limitations in preparing supervisors for work in AI environments. Supervisors in AI-based systems may experience a learning gap, a situation in which they have never performed the task before the adoption of AI (Dohotaru et al., 2025). This gap constrains supervisors' ability to troubleshoot problems and innovate within AI systems.

Organisations should invest in extensive training programmes that enhance supervisory skills for managing AI technologies and foster an environment that values continuous learning (Billiot, 2023). The supervisory frameworks underscore the skills needed for oversight, improvement, empowerment, and harmonisation to promote the responsible and effective use of AI systems. Postgraduate supervision requires clarity in knowledge and the development of supervisory skills as AI takes centre stage in research. Targeted professional development and efforts to bridge learning gaps by educational institutions can better prepare supervisors for work in a changing academic setting, ultimately improving the quality of postgraduate education with AI.

### 1.1 Problem statement

The growing integration of AI technologies into research processes requires urgent analysis of their effects on supervisory relationships, supervisor skills requirements, and the quality of postgraduate education. According to current literature, multiple critical issues have been identified that emphasise this problem (Bolanos et al., 2024). The use of AI tools delivers substantial benefits through increased efficiency and support, but it also raises ethical issues that supervisors must handle carefully to guarantee responsible and effective use (Chauke et al., 2024; Cowling et al., 2023; Khalifa & Albadawy, 2024). Students risk losing critical thinking skills and originality because their dependence on AI-generated content prevents them from conducting thorough research (Kaitharath et al., 2024; Shah & Asad, 2024; Zhai et al., 2024). This reliance on AI prompts doubts about how supervisors can both encourage independent thinking and maintain academic honesty. The literature stresses that supervisors need to develop appropriate skills to effectively guide students in using AI responsibly (Sibiya & Mahosi, 2025). Supervisors must establish explicit guidelines for the use of AI tools so that these technologies support academic work rather than replace student effort (Chiu, 2024; Qutieshat, 2025; Ratnam et al., 2023; Walter, 2024). They need to reassess their conventional supervision skills to integrate AI literacy into their mentorship approaches. Postgraduate supervision requires that supervisors balance the use of AI technology while maintaining ethical standards. Supervisors serve as essential guides for students in evaluating AI-generated materials while helping them avoid common misuse (Alzubi et al., 2025; Li, 2024; Vetter et al., 2024; Wang et al., 2024).

The necessity for supervisors to handle new challenges effectively makes it essential to establish specialised training programmes that equip them with the requisite skills. Educational institutions can enhance supervisory preparation for postgraduate students by developing skills to manage the complexities introduced by AI within a supportive and ethical learning environment. AI is not only revolutionising how postgraduate research is conducted but also emerging as a potential issue that needs to be addressed in terms of supervision and integrity. If not properly guided, its application might have significant implications for the postgraduate research community. In the absence of intervention, AI could lead to an increased reliance on technology by researchers, compromising originality and critical thinking. This dependence, if not addressed, could result in a generation of scholars who prioritise convenience over intellectual rigor, potentially diminishing the overall quality of research. Furthermore, the integration of AI poses a challenge to the role of supervisors, as their traditional mentoring approaches may need to evolve to incorporate AI literacy and ethical considerations. Supervisors who fail to adapt to these changes may find themselves ill-equipped to guide students effectively. Ultimately, the unchecked use of AI might erode trust in postgraduate education, as institutions that overlook the importance of addressing these challenges may risk producing graduates who lack both intellectual independence and a strong moral compass.

The study sought to answer the following question: *What pedagogical mentorship skills are essential for supervisors to guide postgraduate students in the ethical use of AI tools in academic research development?*

## 2. Methodology

A qualitative research method was used to explore supervisors' views on the necessary skills for postgraduate supervision within AI-enhanced settings. The researcher applied an interpretive paradigm to gain thorough contextual insight into supervisors' experiences with AI tool integration (Elbardan et al., 2017; Elliott & Timulak, 2005; Price & Smith, 2021). The study seeks to clarify the intricate supervisory skills and competencies required to supervise postgraduate students who use AI tools by examining participants' experiences. The qualitative research method emphasises understanding the creation of meaning alongside contextual interpretation (Hatch, 2023; Lim, 2024; Tisdell et al., 2025; Tracy, 2024). The researcher selected qualitative approaches because they generate comprehensive and in-depth understandings of participants' experiences and perspectives (Lim, 2024; Moser & Korstjens, 2018; Rosenthal, 2016). This method was utilised to identify subtle elements of supervision practices that quantitative research methods might miss. The study adopts the interpretive research paradigm, which aims to deliver a detailed understanding of supervisors' experiences and facilitates an in-depth examination of the contextual elements affecting supervision skills within AI-enhanced environments. It recognises that supervisors construct their experiences and meanings through social interaction and personal interpretation (Pervin & Mokhtar, 2022; Schwandt, 1994).

### 2.1 Research design

The study employed exploratory phenomenology, chosen for its aim to describe participants' lived experiences and the meanings they attribute to those experiences in naturalistic settings. This exploratory design provides a robust framework for uncovering detailed perspectives on the essential supervision skills required to manage AI-enhanced research environments while upholding academic integrity and student independence amid rapid technological evolution (Eppich et al., 2019; Naz et al., 2022; Ruslin et al., 2022). The phenomenological approach emphasises the essence of lived experience, highlighting the personal significance and subjective interpretation of one's experiences. This orientation facilitated in-depth and reflective inquiry into supervision, enabling the researcher to explore how supervisors interpret, negotiate, and enact integrity, autonomy, and ethical responsibility in their roles while supporting students in an era of swift technological advancement. The exploratory aspect is particularly suitable for contexts where existing theoretical models may be inadequate, allowing the researcher to contribute new insights and understanding related to AI-integrated supervision.

## 2.2 Sampling strategy

Purposive sampling was used to select ten supervisors with experience in supervising post-graduate students. These supervisors were chosen from three public universities in South Africa. Seven participants were drawn from two comprehensive universities, and three from one university of technology. In purposive sampling, researchers deliberately choose participants based on their relevance to the research goals (Campbell et al., 2020; Etikan et al., 2016). The selected participants contributed direct experience by sharing the skills that supervisors must possess when overseeing students in the realm of AI technologies in academic research. Including supervisors from different universities allowed the researcher to identify shared elements and distinct features in various AI integration scenarios within postgraduate research supervision. The variety of institutional backgrounds among participants demonstrates how organisational culture shapes supervision techniques during technological advancement. Analysing supervisors at different points in their careers provided insights into how their perspectives evolve over time amid rapid technological development. The researcher selected participants who met defined experience criteria in AI-enhanced environments to gather data that reveals detailed insights into real-world supervisory challenges. Purposive sampling facilitates a nuanced understanding by examining contextual elements such as disciplinary standards and institutional practices (Campbell et al., 2020; Robinson, 2024) that influence supervisors' methods in implementing new technologies, including generative models.

## 2.3 Data collection

Individual semi-structured interviews were conducted to collect data. These interviews offer a flexible research structure that allows participants to express their thoughts and experiences without being constrained by predefined limits (Iyamu, 2018; Karunarathna et al., 2024; Knott et al., 2022). Qualitative research produces detailed descriptions that reveal subtle aspects (Corbin & Strauss, 2014; Strauss & Corbin, 1998). Narrative-driven themes contribute to the development of theories regarding supervision skills in evolving tech environments, enabling the researcher to explore complex issues in modern postgraduate supervision.

## 2.4 Data analysis

The research applied thematic analysis to explore themes related to supervision skills by analysing semi-structured interview data. Transcripts underwent multiple examinations to detect initial patterns and impressions (Clarke & Braun, 2017; Lochmiller, 2021). The researcher utilised open coding to assign initial labels to sections of data, which were then organised based on their conceptual meanings before being compared to the full dataset (Clarke & Braun, 2017; Kiger & Varpio, 2020; Skjott et al., 2019). Thematic analysis operates within an interpretivist framework to explore the subjective experiences of participants in postgraduate supervision through AI tools (Braun & Clarke, 2023; Ozuem et al., 2022). By following this structured method, researchers can uncover recurring patterns across various experiences while maintaining the ability to develop data through iterative processes (Braun & Clarke, 2022; Peel, 2020). This research investigates the crucial supervision skills needed in technologically advanced academic environments.

## 3. Presentation of Results and Discussion of Findings

This section included the presentation of the data and the discussion of the study's findings. The data presentation was categorised into themes and sub-themes. Three major themes emerged from the data, each with two sub-themes. Data were collected through semi-structured interviews with six women supervisors and four men, who were selected from two comprehensive universities and a university of technology in South Africa. One supervisor is a professor, three hold PhDs, and six supervisors have Master's degrees.

### 3.1 Critical Evaluation Skills

#### *3.1.1 Subtheme 1: Detecting AI overreliance and limitations*

The findings indicate that supervisors should develop a keen ability to discern when students are overly reliant on AI tools in their work. This involves understanding the capabilities and limitations of AI, as well as having a deep insight into the skills and abilities of their students. By reading between the lines, supervisors can identify instances where AI may be doing more than just assisting, potentially undermining the learning process or academic integrity. Participant four from comprehensive university one stated: *"Supervisors should be able to read between the lines to identify when AI is using their students especially knowing the ability of their supervisees"*. Participant one from the university of technology stated that when students struggle to provide accurate sources or authors for the information they present, it often indicates that they have relied heavily on AI without properly understanding or attributing the original sources. "*I check through in-text referencing, usually students who use AI improperly struggle to provide sources/authors, they cannot acknowledge the sources where they got the knowledge […]*".

#### *3.1.2 Subtheme 2: Guiding processes and validation methods*

From the outset, supervisors should guide students through the process of searching for relevant literature, which is a foundational step in academic research. Participant seven from comprehensive university two stated: *"From the beginning when guiding the student is to searching the literature and so forth and I think I've when I reach my students work, I can see this one where is this student reading outdated literature"*. Supervisors must have a comprehensive approach to ensuring that students understand and critically evaluate their research, especially when using AI tools. By requiring students to document their research process and engage in in-depth discussions, the supervisor aims to assess not only the final product but also the thought process and methodology behind it. Participant five from comprehensive university one shared the approach they use to address several key aspects: "*I require students to document their research process, including the prompts they use and how they validate AI-generated information. I also engage in in-depth discussions to* *assess their comprehension of their work*" Supervisors must have a method that they use to detect improper use of AI in student work, focusing on in-text referencing and source acknowledgment.

The research data reveals the essential skill that supervisors possess to detect and handle AI tool misuse within academic settings. These findings align with those of Meinokat and Wagner (2022), which state that supervisors' ability to monitor and understand student behaviour enables them to identify potential digital tool misuse in academic contexts. Furthermore, the findings indicate that supervisors must develop strong detection skills to recognise excessive student reliance on AI by learning about AI capabilities and limitations while gaining an in-depth understanding of their students' abilities. The capacity to read between the lines is a crucial skill that supervisors use to identify situations where AI applications interfere with educational outcomes or academic standards. The findings emphasise the importance of recognising when AI becomes the dominant force in student interactions rather than serving the students. Sibiya and Mahosi (2025) discovered that effective supervision plays a critical role in shaping student writing abilities, as it helps students develop the skills necessary to critically assess literature.

The findings indicate that supervisors should begin by teaching students basic research procedures, including how to find relevant literature. Through this guidance, supervisors can identify emerging problems, such as the use of outdated literature, which suggests insufficient critical analysis of the content. Guiding students through literature research enables supervisors to determine whether students are critically analysing their sources or relying on AI-generated content. Supervisors employ targeted techniques to identify improper AI use by examining in-text citations and source acknowledgements. The inability of students to provide precise source details or author names demonstrates their excessive dependence on AI tools without understanding or attributing original sources.

Students who fail to correctly use AI struggle with citing sources or acknowledging the origins of their information. This highlights the necessity of preserving academic integrity so that students can demonstrate their unique critical thinking and personal effort in their work. Supervisors implement a comprehensive approach to ensure that students conduct critical evaluations of their research work. Students must keep records of their research activities and engage in detailed discussions to evaluate both their final outputs and the reasoning and methods that led to their conclusions. The findings support this approach by requiring students to maintain detailed records of their research process, including the validation of AI-generated data and participation in discussions to assess their understanding. This method enables supervisors to determine whether students use AI to enhance their learning or to substitute their own work with AI assistance. Barkley and Major (2020) noted that student learning assessment methods, such as maintaining research journals and participating in reflective discussions, reveal important insights into students' thinking patterns and academic honesty, allowing tutors to assess authentic student engagement.

### 3.2 Supervisors' pedagogical mentorship skills

#### *3.2.1 Subtheme 1: Essential skills and proactive training*

The supervisors must possess a combination of pedagogical skills to effectively guide students in a digital and technologically advanced academic environment. Participant six from the comprehensive university one outlined these skills: "*Supervisors need digital literacy, critical thinking, ethical awareness, and adaptability*". The supervisors must train students in the responsible use of AI tools. By acknowledging that AI is an integral part of modern academic work, the supervisor emphasises the importance of teaching students how to use AI effectively and ethically. Participant two from the comprehensive university one stated: "*As such, I have to train them how to use it so that when using it they can use it appropriately […]*". The supervisors must guide students on how they can maximise the benefits of AI tools in their academic work. Participant seven from the Comprehensive University Two indicated that by providing guidance on effective AI use, supervisors can help students leverage technology to enhance their research, writing, and critical thinking skills. "*Where possible, supervisors should guide their students on how to use AI to make the best out of it*".

#### *3.2.2 Subtheme 2: Guidance strategies and integrity measures*

Supervisors take a proactive approach to guiding students in their use of AI tools, emphasising the importance of critical thinking and scepticism when relying on AI-generated content. By discouraging full dependence on AI and highlighting the potential inaccuracies or lack of authenticity in AI-generated content, the supervisor aims to foster a balanced and responsible use of technology in academic work. Participant three from Comprehensive University Two stated: "*I totally discourage them from being fully dependent on AI. I even tell them that not all content generated by AI is true or authentic.*" The supervisor guides research students in developing their research focus, emphasising the importance of personal reflection and literature review. By initially focusing on a personal reflection of leadership challenges, the supervisor aims to help students identify a specific research gap related to their leadership experiences. This process is designed to ensure that students engage deeply with their topic and are less likely to rely on AI tools to bypass the reflective process. Participant eight from Comprehensive University Two shared how they help students to keep their research work original: "*So as a first step, I request my [...] students to write a reflection of their leadership challenge or gap; we work on that reflection until it reflects one specific challenge that relates to their leadership, after which I tell the student that they must now go and look for literature [...].*" The supervisor manages plagiarism and ensures academic integrity in student work. A participant outlined that they allow students to write freely and then require them to generate a similarity index report; by doing so, the supervisor aims to educate students about the importance of original work while also providing them with tools to assess their own writing for potential plagiarism. Participant one from the University of Technology shared: "*Hence, I allow them to write freely; thereafter, they proceed to generate a similarity index report. Meanwhile, before this time, they* *are made to know that plagiarism is a grievous offence.*" Supervisors must be able to detect improper use of AI in student work, combining both qualitative analysis and technological tools. Participant ten from Comprehensive University Two shared that they examine writing patterns, check for inconsistencies in argumentation, and utilise AI detection software; the supervisor aims to ensure academic integrity and identify potential misuse of AI tools: "*To detect improper use of AI, I analyse writing patterns, check for inconsistencies in argumentation, and use AI detection software where necessary.*"

The findings show that mentorship in teaching methods is essential for student use of AI tools and academic honesty. Navigating AI implementation in educational settings requires supervisors to demonstrate digital skills, critical thinking abilities, ethical understanding, and flexibility. According to the findings, these essential skills highlight the necessity for supervisors to guide students effectively in the responsible use of AI tools. Supervisors must possess technological pedagogical content knowledge to successfully integrate AI and other technologies into teaching, as this knowledge encompasses technical skills, pedagogical understanding, content expertise, and awareness of ethical technology use, as stated by Yue et al. (2024). Bansal (2023) argues that today's supervisors must teach students to use technology, such as AI, effectively to enhance their learning and academic performance, while also preparing them to thrive in technology-rich environments. The findings reveal that supervisors have the responsibility to teach students to use AI tools properly and effectively, as AI has become essential to contemporary academic work. AI will remain a permanent fixture, so students need to learn to utilise it appropriately. Students learn to apply theoretical principles in practical situations through clear expectation setting and practical training opportunities. Supervisors enable students to unlock AI's full capabilities, which enhances their research skills, writing abilities, and critical thinking.

The findings indicate that supervisors must guide students to use AI tools effectively for maximum benefit while preventing complete reliance on these technologies. According to Robert and Stanworth (2024), educators should instruct students in using digital technologies, including AI, in a way that promotes critical thinking, discernment, and agency, rather than merely relying on these tools, to foster informed and independent learning. The findings emphasise the need for critical evaluation and scepticism when using AI-generated content, as not all information produced by AI systems is factual or genuine. This balanced approach promotes ethical technology use while giving students authority over their research processes and analytical work. Educating students enables them to learn ethical standards and recognise the limitations of AI detection tools while emphasising the integrity of source acknowledgment. Supervisors guide research students in developing research focus through personal reflection and literature review, enabling them to engage deeply with their topics and preventing them from using AI to bypass the reflective process. Ogwueleka (2025) asserts that teaching students about academic integrity, proper citation methods, and the limitations of plagiarism detection will help them understand ethical standards and responsible digital technology use, including AI-generated content. Supervisors uphold academic integrity by educating students about original work while managing plagiarism. They emphasise AI as a supplementary tool to human thought processes, enabling students to utilise AI effectively while preserving personal authorship and genuine work in their academic pursuits. Students should use academic sources to cross-check AI-generated insights to enhance their critical thinking abilities. Educational mentorship is essential for guiding students to responsibly apply AI tools and uphold academic integrity throughout their studies.

**3.3 Adaptive communication skills**

### *3.3.1 Subtheme 1: Proactive guidance and resource support*

Supervisors must provide clear guidance to students on the appropriate use of AI tools in academic work. A supervisor advised that by actively addressing AI use rather than remaining silent, supervisors can ensure that students understand how to leverage AI responsibly and maintain academic integrity. Participant two from Comprehensive University One stated: "*I think also that supervisors must give proper guidance to students on the use of AI rather than keeping silent about the use of AI while students are using* them." Another supervisor supports postgraduate students in their use of AI tools by combining both personal guidance and institutional resources. By encouraging students to attend university-organised training sessions and sharing relevant materials with those who cannot attend, the supervisor ensures that students have access to a wide range of educational opportunities. Participant six from Comprehensive University One shared: "*I also encourage them to attend the AI training that is organised by the university and those who could not attend I make sure I share the slides and the information [...].*" Supervisors take a proactive approach to guide students on integrating technology into their academic work, particularly in tasks such as transcription, mind mapping, and writing enhancement. By encouraging students to use AI tools like transcription software and Grammarly, the supervisor aims to streamline their workflow, improve the quality of their work, and enhance their overall academic experience. Participant nine from the University of Technology echoed: "*I am encouraging my students to use AI tools when they transcribe their audio and video recordings from interviews or observations. [...]. I also encourage my students to use Grammarly to enhance their sentence construction and grammar [...].*"

### *3.3.2 Subtheme 2: Ethical use and authenticity emphasis*

The supervisor emphasises the responsible use of AI as an aid to research rather than a replacement for critical thinking and originality. By providing guidance on ethical practices, proper citation, and validation of AI outputs, the supervisor ensures that students maintain academic integrity while leveraging the benefits of AI. Participant ten from comprehensive university two stated: "*I encourage students to use AI tools responsibly as research aids rather than substitutes for critical thinking. I provide guidance on ethical considerations, proper citation of AI-generated content, and distinguishing between AI-assisted work and original intellectual contributions. I also emphasise the need to* *validate AI-generated insights against peer-reviewed literature."* Supervisors must commit themselves to ensuring that students engage authentically with their academic work, particularly in writing and research. Participant three from the University of Technology shared that they discourage the independent use of AI tools for writing research content; the supervisor emphasises the importance of maintaining academic integrity and sincerity in students' work. *"In terms of using AI tools to assist students with writing [...] their research, I do not allow them to use any tools independently. Like I said, I push them to engage with their work [...], because I feel that it is unethical; it takes away the sincerity of writing."* Participant four from Comprehensive University One encourages students to incorporate their own voice and understanding into their work; the supervisor aims to maintain the authenticity and originality of their research. *"I just tell my students that we need their voice to be in the project. [...]. I try to explain why we are doing the research, as it is not just AI that has to answer the research questions."*

The research findings demonstrate that adaptive communication methods are essential for providing personalised guidance to students based on their specific abilities and research requirements during the use of AI in academic settings. Supervisors provide vital direction regarding appropriate AI tool usage, ensuring that students maintain academic integrity while using AI technologies. Supervisors should engage in discussions about AI use instead of remaining silent to ensure that students learn about responsible AI practices. Postgraduate students receive support from supervisors through both individual mentorship and institutional resources. The findings illustrate this method, where supervisors motivate students to participate in university training sessions and distribute necessary materials to those who cannot attend. Access to diverse educational opportunities allows students to master ethical and effective AI usage. The 2024 study by Ellikkal and Rajamohan shows that supervisors who provide personalised guidance and adaptive communication enhance students' academic achievements and responsible digital technology use, including AI, by customising support to match each student's unique needs and abilities. Supervisors assist students with their academic work by addressing challenges such as outdated literature and restricted internet access.

The findings emphasise that supervisors need to recognise these challenges and provide necessary support by downloading and sharing reading materials to help students access essential resources for success. Supervisors actively help students incorporate technological solutions into their work, promoting the use of AI tools for activities such as transcription and writing improvement. The findings reveal how transcription software and Grammarly contribute to more efficient workflows while enhancing work quality. Effective supervisors identify students' various struggles and deliver specific assistance to tackle these difficulties, leading to better academic performance and encouraging the use of AI tools to improve student work. Supervisors emphasise that AI should support research activities but must not substitute for the essential processes of critical thinking and original creation.

Supervisors instruct students on maintaining academic integrity through ethical practices and proper citation while validating AI outputs. They advise against students using AI tools alone for research content creation, as these tools may undermine genuine student work. The use of independent AI for writing is unethical because it diminishes the sincerity of the work. Supervisors guide students to integrate their personal insights and voice into their research projects, helping to preserve their work's authenticity and originality. Students need to embed their voices into their projects while using AI tools to enhance human insight. The research findings indicate that adaptive communication plays a crucial role in helping students use AI responsibly while preserving their academic integrity and originality. According to Devaki (2025), supervisors must focus on academic integrity in student work while teaching the responsible and critical use of AI tools to ensure that technology supports rather than supplants human insight and original thinking.

## 4. Conclusions and recommendations

The study concluded that supervisors have been positioned as crucial gatekeepers who must navigate the interplay between technology and human intuition, ensuring that AI support does not override student autonomy. This places them in a pivotal role as protectors of the student experience, with mentorship serving as a critical safeguard against over-reliance on AI. The presence of distinct university types (comprehensive, traditional, technology-driven) and the strategies they employ suggest that contextual factors significantly influence how AI tools are adopted and integrated. This implies that a one-size-fits-all policy might not adequately address the diverse realities and resource gaps across different institutions, particularly in the South African higher education context. The convergence of all skills around the principle of academic integrity suggests that effective supervision requires a proactive and comprehensive framing of AI issues, from detection to training. This indicates a need to transform potential vulnerabilities into strengths, effectively preparing scholars for ethical leadership in an increasingly AI-influenced academic environment. The overarching inference from the entire study suggests a shift in the landscape of postgraduate supervision, where AI not only changes the tools available but also potentially alters the epistemology of supervision. This necessitates an ongoing effort in supervisor upskilling to model the behaviours and skills necessary for critical digital citizenship, which in turn shapes graduates who are equipped to ethically navigate AI in educational leadership and broader contexts.

Some of the major recommendations are presented:

- Supervisors need to acquire advanced evaluation skills that will help them identify excessive AI dependency while ensuring students display independent thought and original ideas.
- Supervisors should actively guide students through core research methodologies, including exhaustive literature reviews and correct source citation, to maintain academic standards.

- Supervisors must evaluate students' understanding and promote ethical AI practices by requiring documentation of the research process along with thorough discussion involvement.
- The effective supervision of contemporary postgraduate education relies on finding a balance between using AI tools and fostering independent critical thinking and ethical standards.
- Specialised training programmes need to be developed by educational institutions to enhance supervisors' understanding of AI systems and their ability to assess critically.
- Supervisors require training programmes that equip them with the necessary tools and strategies to guide students on the ethical and responsible use of AI technologies.
- Strict rules must be established regarding how AI should be utilised within postgraduate research initiatives.
- Research should focus on creating scalable supervisor training frameworks that meet the unique demands of AI in postgraduate supervision across different academic disciplines and institutional environments.

## 5. Declarations

**Funding:** This research did not receive any external funding.

**Conflicts of Interest:** The author declares no conflict of interest.

**Use of Artificial Intelligence:** The current work was created without the assistance of artificial intelligence technologies, as confirmed by the author.

## References

Ahrweiler, P., Späth, E., Siqueiros García, J. M., Capellas, B. L., & Wurster, D. (2025). Inclusive technology co-design for participatory AI. In *Participatory artificial intelligence in public social services: From bias to fairness in assessing beneficiaries* (pp. 35-62). https://doi.org/10.1007/978-3-031-71678-2_2

Alzubi, A. A. F., Nazim, M., & Alyami, N. (2025). Do AI-generative tools kill or nurture creativity in EFL teaching and learning? *Education and Information Technologies*, 1-38. https://doi.org/10.1007/s10639-025-13409-8

Amer, M. L. B., Rababah, L. M., & Rababah, M. A. (2025). Exploring ChatGPT as a tool for thesis writing: Perspectives of EFL supervisors in Jordanian universities. *Journal of Language Teaching and Research*, *16*(1), 33-42. https://doi.org/10.17507/jltr.1601.04

Bainey, K. (2024). *AI-driven project management: Harnessing the power of artificial intelligence and ChatGPT to achieve peak productivity and success.* John Wiley & Sons.

Bankins, S., Ocampo, A. C., Marrone, M., Restubog, S. L. D., & Woo, S. E. (2024). A multilevel review of artificial intelligence in organizations: Implications for organizational behavior research and practice. *Journal of Organizational Behavior*, *45*(2), 159-182. https://doi.org/10.1002/job.2735

Bansal, S. K. (2023). Significance and advancement of educational technology: Holistic development of learners in the digital age. *Available at SSRN* 4863716. https://ssrn.com/abstract=4863716 or http://dx.doi.org/10.2139/ssrn.4863716
Barkley, E. F., & Major, C. H. (2020). *Student engagement techniques: A handbook for college faculty*. John Wiley & Sons.
Bhadoriya, E. (2024). Unlocking simulation modeling proficiency: AI as a key to autonomous expertise development. https://ir.library.illinoisstate.edu/etd/2009
Billiot, T. (2023). Continuous learning and advancing technologies: A framework for professional development and training in artificial intelligence. *Development and Learning in Organizations: An International Journal*, *37*(3), 28-31. https://doi.org/10.1108/DLO-04-2022-0064
Bolaños, F., Salatino, A., Osborne, F., et al. (2024). Artificial intelligence for literature reviews: Opportunities and challenges. *Artificial Intelligence Review*, *57*, 259. https://doi.org/10.1007/s10462-024-10902-3
Braun, V., & Clarke, V. (2022). Conceptual and design thinking for thematic analysis. *Qualitative Psychology*, *9*(1), 3. https://doi.org/10.1037/qup0000196
Braun, V., & Clarke, V. (2023). Toward good practice in thematic analysis: Avoiding common problems and becoming a knowing researcher. *International Journal of Transgender Health*, *24*(1), 1-6. https://doi.org/10.1080/26895269.2022.2129597
Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., & Walker, K. (2020). Purposive sampling: Complex or simple? Research case examples. *Journal of Research in Nursing*, *25*(8), 652-661. https://doi.org/10.1177/1744987120927206
Chauke, T. A., Mkhize, T. R., Methi, L., & Dlamini, N. (2024). Postgraduate students' perceptions on the benefits associated with artificial intelligence tools on academic success: In case of ChatGPT AI tool. *Journal of Curriculum Studies Research*, *6*(1), 44-59. https://doi.org/10.46303/jcsr.2024.4
Chiu, T. K. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence*, *6*, 100197. https://doi.org/10.1016/j.caeai.2023.100197
Clarke, V., & Braun, V. (2017). Thematic analysis. *The Journal of Positive Psychology*, *12*(3), 297-298. https://doi.org/10.1080/17439760.2016.1262613
Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications.
Corey, G., Haynes, R. H., Moulton, P., & Muratori, M. (2020). *Clinical supervision in the helping professions: A practical*
*guide*. John Wiley & Sons.
Cowling, M., Crawford, J., Allen, K. A., & Wehmeyer, M. (2023). Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research

supervision. *Australasian Journal of Educational Technology*, *39*(4), 89-103. https://doi.org/10.14742/ajet.8598
Davys, A., & Beddoe, L. (2020). *Best practice in professional supervision: A guide for the helping professions*. Jessica Kingsley Publishers.
Devaki, V. (2025). Academic integrity and human cognitive development of learners: Is artificial intelligence replacing the human brain? In *AI applications and strategies in teacher education* (pp. 195-222). IGI Global. http://doi.org/10.4018/979-8-3693-5443-8.ch007
Dohotaru, M., Palta, Y., Prisacaru, M., & Shin, J. H. (2025). *AI for risk-based supervision*. World Bank.
Elbardan, H., Kholeif, A. O., Elbardan, H., & Kholeif, A. O. R. (2017). An interpretive approach for data collection and analysis. In *Enterprise resource planning, corporate governance and internal auditing: An institutional perspective* (pp. 111-165). https://doi.org/10.1007/978-3-319-54990-3_5
Ellikkal, A., & Rajamohan, S. (2024). AI-enabled personalized learning: Empowering management students for improving engagement and academic performance. *Vilakshan-XIMB Journal of Management*. https://doi.org/10.1108/XJM-02-2024-0023
Elliott, R., & Timulak, L. (2005). Descriptive and interpretive approaches to qualitative research. In A handbook of research methods for clinical and health psychology (Vol. 1, pp. 147-159).
Eppich, W. J., Gormley, G. J., & Teunissen, P. W. (2019). In-depth interviews. In Healthcare simulation research: A practical guide (pp. 85-91). https://doi.org/10.1007/978-3-030-26837-4_12
Etikan, I., Musa, S. A., & Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1-4. https://doi.org/10.11648/j.ajtas.20160501.11
George, B., & Wooden, O. (2023). Managing the strategic transformation of higher education through artificial intelligence. *Administrative Sciences*, *13*(9), 196. https://doi.org/10.3390/admsci13090196
Hatch, J. A. (2023). *Doing qualitative research in education settings*. State University of New York Press.
Iyamu, T. (2018). Collecting qualitative data for information systems studies: The reality in practice. *Education and Information Technologies*, *23*, 2249-2264. https://doi.org/10.1007/s10639-018-9718-2
Jarrahi, M. H., Askay, D., Eshraghi, A., & Smith, P. (2023). Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons*, *66*(1), 87-99. https://doi.org/10.1016/j.bushor.2022.03.002
Kaitharath, M. F., Shifna, K. N., & Anuradha, P. S. (2024). Generative AI and its impact on creative thinking abilities in higher education institutions. In *Impacts of generative AI on*

*creativity in higher education* (pp. 141-176). IGI Global. https://doi.org/10.4018/979-8-3693-2418-9.ch006

Karunarathna, I., Gunasena, P., Hapuarachchi, T., & Gunathilake, S. (2024). The crucial role of data collection in research: Techniques, challenges, and best practices. *Uva Clinical Research*, 1-24.

Khalifa, M., & Albadawy, M. (2024). Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update*, 100145. https://doi.org/10.1016/j.cmpbup.2024.100145

Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical Teacher*, *42*(8), 846-854.

Knott, E., Rao, A. H., Summers, K., & Teeger, C. (2022). Interviews in the social sciences. *Nature Reviews Methods Primers*, *2*(1), 73. https://doi.org/10.1038/s43586-022-00150-6

Le, H. (2024). *The evolving role of programmers in an AI-chatbot dominated world: Challenges, adaptation strategies, and future prospects* (Doctoral dissertation, University of the Cumberlands).

Lee, A. (2019). *Successful research supervision: Advising students doing research*. Routledge. https://doi.org/10.4324/9781351234986

Li, Z. (2024). *Generative AI in higher education academic assignments: Policy implications from a systematic review of student and teacher perceptions* (Doctoral dissertation, Massachusetts Institute of Technology). https://rightsstatements.org/page/InC-EDU/1.0/

Lim, W. M. (2024). What is qualitative research? An overview and guidelines. *Australasian Marketing Journal*, 14413582241264619. https://doi.org/10.1177/14413582241264619

Lochmiller, C. R. (2021). Conducting thematic analysis with qualitative data. *The Qualitative Report*, *26*(6), 2029-2044. https://nsuworks.nova.edu/tqr

Mack, O. (2023). Model for professionals to achieve effective supervision of artificial intelligence. MIT Computational Law Report.

Meinokat, P., & Wagner, I. (2022). Causes, prevention, and interventions regarding classroom disruptions in digital teaching: A systematic review. *Education and Information Technologies*, *27*(4), 4657-4684. https://doi.org/10.1007/s10639-021-10795-7

Misra, S., Katz, B., Roberts, P., Carney, M., & Valdivia, I. (2024). Toward a person-environment fit framework for artificial intelligence implementation in the public sector. *Government Information Quarterly*, *41*(3), 101962.

Morandini, S., Fraboni, F., De Angelis, M., Puzzo, G., Giusino, D., & Pietrantoni, L. (2023). The impact of artificial intelligence on workers' skills: Upskilling and reskilling in organisations. *Informing Science*, *26*, 39-68. https://hdl.handle.net/11585/917132

Moser, A., & Korstjens, I. (2018). Series: Practical guidance to qualitative research. Part 3: Sampling, data collection and analysis. *European Journal of General Practice*, *24*(1), 9-18. https://doi.org/10.1080/13814788.2017.1375091

Naz, N., Gulab, F., & Aslam, M. (2022). Development of qualitative semi-structured interview guide for case study research.

Ogwueleka, F. N. (2025). Plagiarism detection in the age of artificial intelligence: Current technologies and future directions. *AI and Ethics, Academic Integrity and the Future of Quality Assurance in Higher Education*, 10.

Ozuem, W., Willis, M., & Howell, K. (2022). Thematic analysis without paradox: Sensemaking and context. *Qualitative Market Research: An International Journal, 25(*1), 143-157. https://doi.org/10.1108/QMR-07-2021-0092

Peel, K. L. (2020). A beginner's guide to applied educational research using thematic analysis. *Practical Assessment Research and Evaluation, 25*(1). https://doi.org/10.7275/ryr5-k983

Pervin, N., & Mokhtar, M. (2022). The interpretivist research paradigm: A subjective notion of a social context. *International Journal of Academic Research in Progressive Education and Development, 11*(2), 419–428. http://dx.doi.org/10.6007/IJARPED/v11-i2/12938

Price, H. E., & Smith, C. (2021). Procedures for reliable cultural model analysis using semi-structured interviews. *Field Methods, 33*(2), 185–201. https://doi.org/10.1177/1525822X20982725

Qutieshat, A. (2025). *Artificial intelligence in higher education: A contemporary examination of Illich's theories*. Springer Nature.

Rajabi, P. (2023). Investigating computing science students' and educators' initial perceptions of ChatGPT in higher education.

Ratnam, M., Sharma, B., & Tomer, A. (2023). ChatGPT: Educational artificial intelligence. *International Journal of Advanced Trends in Computer Science and Engineering, 12*(2). https://doi.org/10.30534/ijatcse/2023/091222023

Robert, M., & Stanworth, M. (2024). Shaping future education: Pedagogical strategies for integrating critical thinking and AI.

Robinson, R. S. (2024). Purposive sampling. In *Encyclopedia of quality of life and well-being research* (pp. 5645-5647). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-17299-1_2337

Rosenthal, M. (2016). Qualitative research methods: Why, when, and how to conduct interviews and focus groups in pharmacy research. *Currents in Pharmacy Teaching and Learning, 8(4), 509-516. https://doi.org/10.1016/j.cptl.2016.03.021*

Ruslin, R., Mashuri, S., Rasak, M. S. A., Alhabsyi, F., & Syam, H. (2022). Semi-structured interview: A methodological reflection on the development of a qualitative research instrument in educational studies. *IOSR Journal of Research & Method in Education (IOSR-JRME)*, *12*(1), 22–29. https://doi.org/10.9790/7388-1201052229

Schwandt, T. A. (1994). Constructivist, interpretivist approaches to human inquiry. In *Handbook of qualitative research* (pp. 118–137).

Shah, S. S., & Asad, M. M. (2024). Impact of critical thinking approach on learners' dependence on innovative transformation through artificial intelligence. In *The evolution of artificial intelligence in higher education: Challenges, risks, and ethical considerations* (pp. 161–182). Emerald Publishing Limited. https://doi.org/10.1108/978-1-83549-486-820241010

Sibiya, A., & Mahosi, B. N. (2025). The role of supervisors in developing postgraduate literacies. https://doi.org/10.20944/preprints202501.1321.v1

Skjott Linneberg, M., & Korsgaard, S. (2019). Coding qualitative data: A synthesis guiding the novice. *Qualitative Research Journal*, *19*(3), 259–270. https://doi.org/10.1108/QRJ-12-2018-0012

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory*.

Tisdell, E. J., Merriam, S. B., & Stuckey-Peyrot, H. L. (2025). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.

Tracy, S. J. (2024). *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact*. John Wiley & Sons.

Vetter, M. A., Lucia, B., Jiang, J., & Othman, M. (2024). Towards a framework for local interrogation of AI ethics: A case study on text generators, academic integrity, and composing with ChatGPT. *Computers and Composition*, *71*, 102831. https://doi.org/10.1016/j.compcom.2024.102831

Walter, Y. (2024). Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, *21*(1), 15. https://doi.org/10.1186/s41239-024-00448-3

Wang, H., Dang, A., Wu, Z., & Mac, S. (2024). Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines. *Computers and Education: Artificial Intelligence*, *7*, 100326. https://doi.org/10.1016/j.caeai.2024.100326

Yue, M., Jong, M. S. Y., & Ng, D. T. K. (2024). Understanding K–12 teachers' technological pedagogical content knowledge readiness and attitudes toward artificial intelligence education. *Education and Information Technologies*, 1–32. https://doi.org/10.1007/s10639-024-12621-2

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments*, *11*(1), 28. https://doi.org/10.1186/s40561-024-00316-7

CHAPTER SEVEN

# Challenges and Risks of Utilising AI Tools for Postgraduate Supervision in Higher Education: Perspectives from South Africa

Peter Babajide Oloba[1]

AFFILIATIONS

[1]Department of Languages, Cultural Studies and Applied Linguistics, Faculty of Humanities, University of Johannesburg, Johannesburg, South Africa.

CORRESPONDENCE

Email: olobapeter4u@gmail.com

REFERENCE

Oloba, P. B. (2026). Challenges and risks of utilising AI tools for postgraduate supervision in higher education: Perspectives from South Africa. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 111-129). ERRCD Forum. https://doi.org/10.38140/obp4-2026-07



**Abstract:** The integration of artificial intelligence (AI) tools into postgraduate supervision in higher education has accelerated globally, offering opportunities to enhance efficiency in research processes and academic mentoring. However, limited empirical evidence exists regarding the risks and challenges of this integration, particularly within Global South contexts such as South Africa. This study investigates the challenges associated with the use of AI tools in postgraduate supervision from a South African perspective. Anchored in a constructivist paradigm, the study adopts a qualitative research design, employing semi-structured interviews with 20 purposively selected participants—10 postgraduate students and 10 supervisors from faculties that are actively integrating AI into supervisory practices. Data were analysed thematically using qualitative content analysis. The findings identify six key challenges: increasing dependence on AI that may erode students' critical thinking and originality; sufficient digital literacy and institutional support; financial and sustainability constraints; the questionable reliability and accuracy of AI-generated outputs; ethical dilemmas and limited cultural contextualisation; and resistance to technological change among supervisors. While acknowledging the potential of AI to enhance research productivity and the quality of supervision, the study cautions against its uncritical adoption, which may compromise academic integrity, creativity, and equity. It recommends institutional strategies, including subsidised AI access, structured training on ethical and critical AI use, the embedding of digital literacy in postgraduate curricula, and the fostering of collaboration with AI developers to ensure culturally relevant systems. A context-sensitive approach is essential to balance the affordances of AI with the preservation of human intellectual agency and critical scholarly engagement in postgraduate supervision.

***Keywords:*** Artificial intelligence, critical thinking, contextual sensitivity, cultural sensitivity, digital literacy, higher education, postgraduate supervision.

## 1. Introduction

The integration of artificial intelligence (AI) tools into higher education has accelerated significantly in recent years, with an increasing adoption in teaching, learning, and research supervision. AI platforms—such as ChatGPT, Grammarly, Elicit, Jasper AI, QuillBot, Scite, Research Rabbit, Consensus, and various research summarisation tools—provide postgraduate students and supervisors with opportunities to streamline administrative tasks, enhance writing quality, and access information efficiently (Gasaymeh et al., 2024; Raheem et al., 2023). In the domain of postgraduate supervision specifically, AI is increasingly employed to assist with literature reviews, methodology design, data analysis, writing support, and even the drafting of research proposals. However, despite the growing reliance on AI tools in postgraduate supervision, there remains an insufficient understanding of the challenges and risks that these technologies introduce within higher education systems, particularly in non-Western contexts such as South Africa.

While an expanding body of global scholarship has investigated the ethical implications, biases, and technical limitations of AI in education (Al-Zahrani & Alasmari, 2024; Akgun & Greenhow, 2022), few studies have focused specifically on the postgraduate supervision environment. Even fewer have examined these issues from a Global South perspective, where socio-cultural, linguistic, infrastructural, and epistemological differences profoundly shape the experiences and impacts of AI use (Ade-Ibijola & Okonkwo, 2023; Roche et al., 2023). In African contexts such as South Africa, postgraduate education often involves navigating indigenous knowledge systems, multilingual realities, and community-based research protocols that AI tools, largely developed in Western, high-income settings, are poorly equipped to address. Consequently, there exists a critical gap in understanding how AI-related challenges manifest differently outside of Euro-American academic frameworks. This gap constitutes the central problem underpinning this study: despite the rapid integration of AI tools into postgraduate supervision, there is a lack of empirical evidence concerning the risks and challenges these technologies pose within South African higher education. The existing AI tools may inadvertently undermine critical thinking, compromise originality, misrepresent culturally grounded knowledge, and exacerbate digital inequalities—issues that remain underexplored in current research (Ade-Ibijola & Okonkwo, 2023; Akgun & Greenhow, 2022). In the absence of a clear, context-specific understanding of these risks, South African universities face significant uncertainty in developing appropriate guidelines, training frameworks, and ethical policies to regulate AI use in supervision.

In light of the rapid evolution of AI technologies and their increasingly unpredictable influence on research practices, it is imperative to critically evaluate their implications for postgraduate supervision in South Africa. This study is, therefore, justified by the pressing necessity to generate evidence-based insights that can inform institutional policy, safeguard academic integrity, and guide the ethical and culturally sensitive integration of AI. As supervisors and students increasingly rely on AI in the absence of robust institutional frameworks, there exists a significant risk that uncritical adoption may undermine human intellectual agency, marginalise indigenous knowledge systems, and exacerbate existing inequities. Addressing this gap is essential to ensure that AI functions as an enabler, rather than a threat, to quality, inclusive, and contextually relevant postgraduate education in South Africa. This study aims to address this underexplored area by focusing on the South African higher education context. In doing so, it seeks to contribute to a more nuanced and situated body of knowledge on AI in education, ensuring that technological innovation supports rather than undermines academic quality, ethical research practices, and local relevance in postgraduate education.

### 1.1 Statement of the problem

The rapid proliferation of AI tools within higher education has begun to reshape postgraduate research and supervision practices, often at a pace that exceeds the development of coherent institutional, ethical, and pedagogical frameworks. In South Africa, postgraduate students and supervisors are increasingly engaging with AI tools to support literature reviews, academic writing, data analysis, and research planning (Al-Zahrani, 2024; Dwivedi et al., 2023). While these technologies promise enhanced efficiency and productivity, their uncritical adoption raises substantive concerns regarding the integrity and quality of postgraduate education. Emerging evidence indicates that excessive reliance on AI tools may undermine critical thinking, scholarly independence, and the formation of academic identity, which are core outcomes of postgraduate training (Gerlich, 2025; Zhai et al., 2024). These risks are particularly pronounced within supervision contexts, where mentorship, intellectual guidance, and epistemic judgement are central, yet increasingly mediated by opaque and poorly understood technological systems.

The problem is further intensified by the South African higher education landscape, which is shaped by persistent socio-economic inequalities, uneven digital infrastructure, multilingual realities, and the importance of indigenous and community-based knowledge systems. AI tools are predominantly developed within Western epistemological and cultural frameworks, limiting their capacity to engage meaningfully with local research contexts, ethical protocols, and culturally embedded forms of knowledge production (Ade-Ibijola & Okonkwo, 2023; Ofosu-Asare, 2024). As a result, the integration of AI risks reinforcing existing inequities, marginalising local and indigenous knowledge, and compromising ethical research practices, particularly where access, affordability, and digital literacy remain uneven (Akgun & Greenhow, 2022; Al-Zahrani, 2024). Despite the accelerating uptake of AI tools in postgraduate supervision, there remains a notable lack of empirical, context-specific evidence to guide institutional policy, supervisory practice, and ethical governance in South African universities. This absence of grounded insight leaves both students and supervisors navigating AI adoption without clear guidance, increasing the likelihood that AI functions as a destabilising rather than enabling force in postgraduate education. Addressing this problem is therefore imperative to ensure that AI supports, rather than erodes, scholarly rigour, ethical integrity, and contextual relevance in postgraduate supervision.

Hence, this study answers the research question: *What are the challenges and risks associated with utilising AI tools for postgraduate supervision in higher education?*

## 2. Methodology

This study was framed within the constructivist paradigm, which emphasises the understanding of individuals' lived experiences and the subjective meanings they attribute to those experiences (Turin et al., 2024). Constructivism presupposes that reality is socially constructed, thus knowledge emerges through interaction between the researcher and participants (Clarke et al., 2023). In the context of this study, a constructivist lens was deemed appropriate for exploring how students and supervisors in South Africa perceive the challenges and risks associated with the utilisation of AI tools in postgraduate supervision. A qualitative research approach was adopted to facilitate an in-depth exploration of participants' perspectives (Lim, 2024). Qualitative research is especially suited for investigating complex phenomena where little is known and where understanding participants' experiences, perceptions, and meanings is paramount (Creswell & Creswell, 2017). Given the nascent nature of AI use in higher education supervision, this approach permitted the capture of rich, detailed descriptions. The study employed a phenomenological research design. Phenomenology concentrates on understanding how individuals experience a specific phenomenon by uncovering commonalities in their lived experiences (Gill, 2020). In this instance, the design enabled the researcher to delve into the lived experiences of postgraduate students and supervisors regarding the challenges and risks encountered when utilising AI tools in postgraduate supervision.

The population for this study comprised postgraduate students and supervisors from ten universities in South Africa. In total, twenty participants engaged in the interview process: ten postgraduate students and ten supervisors. This sample size was deliberately selected, as qualitative phenomenological studies prioritise depth of insight over numerical representativeness; smaller samples facilitate the emergence of rich, detailed accounts of lived experiences (Creswell & Poth, 2016). A sample of twenty was deemed sufficiently large to capture diverse perspectives across different universities while remaining manageable for conducting in-depth, interpretive analysis, which is central to phenomenological inquiry (Vagle, 2018). Furthermore, purposive sampling ensured that all selected participants had direct experience with the integration or navigation of AI tools in postgraduate supervision, rendering them information-rich cases capable of providing meaningful data relevant to the study's aim. Purposive sampling was appropriate, as it allowed the researcher to deliberately select participants possessing direct experience with the phenomenon under investigation, thereby ensuring rich, pertinent insights (Campbell et al., 2020). A descriptive analysis of participants' demographic characteristics was compiled to contextualise their perspectives.

The sample included ten postgraduate students and ten supervisors, representing a range of disciplines across humanities, education, science, law, engineering, commerce, arts, design, and architecture faculties. The student group consisted of honours, master's, and doctoral candidates, while supervisors varied in supervisory experience from early-career academics to senior scholars. Participants were drawn from both historically advantaged and historically disadvantaged universities, thereby ensuring diversity in institutional contexts and enriching the phenomenological insights into AI use in postgraduate supervision. This demographic variation strengthened the study by capturing a wide spectrum of lived experiences and contextual influences. Semi-structured interviews served as the primary method of data collection. This approach provided the flexibility to explore predetermined topics while also allowing participants the freedom to discuss issues they deemed significant (Ruslin et al., 2022). Although interviews were guided by an interview schedule, they permitted probing and follow-up questions to clarify and expand upon participant responses. The data were analysed using qualitative content analysis, which involves a systematic coding and categorisation process to identify patterns, themes, and categories within the data (Zhang & Wildemuth, 2009). Content analysis enabled the researcher to interpret and organise the complex data into meaningful units that addressed the research questions regarding the challenges and risks associated with the use of AI tools for postgraduate supervision.

## 3. Presentation of Results and Discussion of Findings

In response to the research question regarding the challenges and risks associated with utilising AI tools for postgraduate supervision in higher education, six key themes emerged: dependence on AI tools; inadequate literacy, support, and integration of supervision; cost and sustainability; quality and accuracy of AI output; ethical concerns, including cultural and contextual sensitivity; and resistance to change.

### 3.1 Dependence on AI tools

An emerging risk associated with the use of AI tools in postgraduate supervision is the increasing dependence on these technologies, which may compromise students' critical thinking, analytical abilities, and independent academic development. Both supervisors and students shared concerns about how over-reliance on AI is affecting the quality of research training and academic skills. A senior humanities supervisor raised concerns about students' excessive reliance on AI for major sections of their academic work: "*Some students now rely on AI to write entire sections of their research. It is worrying because they are not developing the necessary analytical or writing skills* themselves." Similarly, a master's student from the Education Faculty acknowledged the tendency to overuse AI tools, particularly during tasks like literature reviews: "*AI tools help speed things up, especially when doing literature reviews, but I find myself depending on them too much instead of reading and analysing articles on my* own." An early-career supervisor in the sciences noted the superficial quality of some student submissions, attributing it to the overuse of AI tools: "*I have noticed a trend where students submit work that looks polished but lacks depth. It becomes clear in discussions that the critical engagement is missing, possibly because the AI did most of the* work." A doctoral student in engineering also shared personal struggles with diminished writing confidence after habitual use of AI platforms: "*At first, using AI tools was exciting, but now I struggle to write without them. It is like I have lost confidence in my own ability to start from* scratch." Echoing this concern, a senior supervisor from the Commerce Faculty emphasised the need for proper training on appropriate AI use: "*We need to teach students when and how to use AI appropriately. Right now, some are just outsourcing their thinking to* machines." An honours student in the Arts and Design Faculty expressed doubt about the authenticity of their own academic growth: "*Sometimes I wonder if I am learning or just editing what the AI generates. It is convenient, but I am not sure I am growing as a* researcher."

Supervisors particularly noted that some students are relying heavily on AI tools for composing significant sections of their research, resulting in underdeveloped analytical and academic writing skills. This observation aligns with findings by Gerlich (2025), who cautions that while AI technologies provide efficiency, they can also deskill students by automating cognitive processes that are essential for achieving higher education learning outcomes, such as critical reasoning and academic argumentation. Students' own reflections reveal an internal struggle with dependence, particularly during literature reviews, where AI tools expedite the process but at the expense of deep reading and personal analysis. Zhai et al. (2024) similarly argue that an overreliance on AI for information processing can disincentivise deeper cognitive engagement, thereby undermining one of the core purposes of postgraduate education: developing the ability to critically interrogate texts and construct original arguments. The concern regarding the superficial quality of submissions—polished work lacking critical depth—is particularly alarming. Supervisors observed that although AI-generated text may appear well-written, it often conceals the absence of genuine analytical thought, which becomes evident during oral discussions. This echoes the critique by Delcker et al. (2024), who highlight the veneer of competence that AI can create in education, wherein surface-level outputs are mistakenly recognised as indicative of deep learning.

Moreover, students reported a loss of confidence in their own writing abilities after frequent use of AI, feeling unable to initiate academic writing without machine assistance. This aligns with Gerlich (2025), who warns that habitual use of AI can lead to academic learned helplessness, whereby students doubt their own capabilities and become overly reliant on AI-generated content to commence or structure their work. Supervisors emphasised the necessity for explicit training on the appropriate and ethical use of AI, as well as guidance on when and how to integrate AI tools without outsourcing one's intellectual labour. Without clear guidelines, students risk not only stunting their academic growth but also breaching ethical boundaries in their research processes. Additionally, students' doubts about the authenticity of their learning, questioning whether they are genuinely developing as researchers, point to a deeper existential risk: the erosion of academic identity. If postgraduate education fundamentally aims to develop scholars capable of independent thought and the contribution of original knowledge, then unchecked dependence on AI threatens the very core of that mission (Zhai et al., 2024).

### 3.2 Inadequate literacy, support, and supervisory integration

Participants indicate that one of the challenges associated with the use of AI tools for postgraduate supervision is the lack of adequate digital literacy, institutional support, and the seamless integration of these technologies into existing supervisory practices. One of the Honours students in the faculty of Humanities shared their difficulty in navigating AI platforms without formal guidance: "*I often hear about ChatGPT and other AI platforms, but honestly, I do not know how to use them in a way that supports my research. There is no proper orientation or guidance.*" A master's student in engineering highlighted the knowledge gap that exists among supervisors, which affects the effective use of AI tools in the supervision process: "*Most supervisors in my department are not familiar with AI tools, so even if we wanted to use them, there is a knowledge gap that affects how we are supervised.*" A doctoral student in Education recounted the lack of structured training and the need for self-directed learning: "*We were introduced to AI tools without any structured training. I had to teach myself through YouTube videos. It is not sustainable or effective.*" Similarly, a senior supervisor in the Faculty of Science described the overwhelming pressure of incorporating AI tools without proper training or institutional preparation: "*I am expected to incorporate AI tools into my supervision, but I have received no formal training. It becomes overwhelming and time-consuming, especially with an already heavy workload.*" Another early-career engineering supervisor pointed to the absence of technical support structures within the institution: "*There is no institutional support or technical team to help us when we get stuck using these tools. You either figure it out yourself or abandon it.*" Expressing frustration over the lack of clear frameworks, another senior supervisor in Arts and Design stated, "AI *tools are supposed to help streamline supervision, but in reality, they have added extra layers of confusion because we don't have a clear framework on how to use them in our workflow.*" A master's student in law also reflected on how supervisors' lack of confidence with AI tools limits collaboration possibilities: "*My supervisor is not very confident using AI tools, so we just stick to emails and WhatsApp. It is frustrating because I know there are better ways to collaborate.*" A supervisor in Humanities stressed the need for better integration of AI tools into traditional supervisory models: "*We need proper integration of AI tools into our existing supervision models. Right now, it feels like we are trying to fit a square peg into a round hole.*"

The findings indicate that both students and supervisors emphasised how a lack of structured training, technical support, and clear implementation frameworks undermines the potential benefits of AI technologies. The issue of digital literacy was strongly highlighted. Several students reported difficulties in navigating AI platforms such as ChatGPT without formal guidance. One participant noted reliance on self-directed learning through informal sources, such as YouTube videos. This reflects broader concerns raised by Filiz et al. (2025), who assert that the successful adoption of AI in education heavily depends on the digital competencies of both students and educators. In the absence of structured training, students are often left to navigate these technologies independently, resulting in inconsistent usage and superficial engagement with AI tools (Zhou et al., 2024). The knowledge gaps among supervisors further exacerbate this challenge. Students observed that supervisors' unfamiliarity with AI tools adversely affects the quality of supervision they receive. This aligns with Al-Zahrani (2024), who contends that if supervisors are not adequately trained or confident in using AI, they are unlikely to encourage or model effective usage for their students. Rather than serving as a collaborative tool to enhance supervision, AI becomes a source of disconnection.

The data also revealed that the absence of clear frameworks for AI integration generates confusion instead of streamlining the supervisory process. Supervisors reported that AI tools often feel like an additional burden rather than an enabler of efficient supervision, as they are introduced without alignment to existing academic workflows. Southworth et al. (2023) argue that AI tools must be integrated thoughtfully into existing pedagogical models, rather than being deployed as add-ons or afterthoughts. When integration is poorly planned, it increases cognitive overload for both supervisors and students, thereby reducing the effectiveness of supervision. Furthermore, the lack of confidence among supervisors has downstream effects on collaboration. One student noted that, due to their supervisor's discomfort with AI tools, communication reverted to traditional methods such as emails and WhatsApp, which limited opportunities for AI-supported collaboration. As highlighted by Haleem et al. (2017), digital collaboration tools can enhance feedback loops, peer interaction, and mentorship, but only when users possess the necessary confidence and competence in their use. Ultimately, the call for improved integration of AI tools into existing supervisory models reflects the necessity for systematic change. Simply introducing AI tools without re-evaluating supervision models results in a scenario where the technology appears mismatched and ineffective.

### 3.3 Cost and sustainability

Cost and sustainability emerged as significant challenges associated with the utilisation of AI tools for postgraduate supervision in higher education. An honours student emphasised the unaffordability of AI tools without institutional support: "*Some of these AI tools come with monthly subscription costs that are simply not feasible for students like us. If the university does not cover them, we can't afford to use them consistently*." Similarly, an early-career supervisor raised concerns about the sustainability of AI platforms once initial funding or trial periods end: "*We were introduced to a promising AI platform for supervision, but the institution couldn't maintain the subscription after the trial period. It raises concerns about how sustainable these tools really are for long-term use*." A master's student highlighted infrastructural barriers, particularly for those in less-resourced areas: "*AI tools are great in theory, but without stable internet and access to high-performance devices, many of us in rural areas are left behind*." A mid-senior supervisor also pointed to the financial burden institutions face in maintaining AI tools beyond mere initial access: "*The cost of maintaining and updating AI platforms can be burdensome for our institution. It is not just about access but about sustaining that access over time*." A doctoral candidate noted how financial challenges at the student level can lead to exclusion, undermining the goal of inclusive supervision: "*I have seen students drop out of AI-supported initiatives simply because they could not keep up with the cost of data or did not have devices compatible with the tools. It defeats the purpose of inclusive supervision*."

The data presentation underscores cost and sustainability as significant challenges to the effective utilisation of AI tools in postgraduate supervision. Both students and supervisors expressed concerns regarding subscription fees, infrastructure requirements, and the long-term financial burden associated with the integration of AI in higher education contexts. Financial accessibility emerged as a major barrier. Students reported that many AI tools require monthly subscription fees that are unaffordable without institutional support. This finding aligns with Sabiteka et al. (2025), who indicated that affordability is one of the key barriers to the adoption of educational technology, particularly in low- and middle-income contexts. Without subsidised access, AI tools risk exacerbating inequalities among students, particularly those from disadvantaged backgrounds. As one student noted, unless universities cover these costs, consistent use of AI tools remains unattainable for many. Furthermore, the sustainability of AI initiatives beyond initial trials was another concern. A supervisor pointed out that while institutions often introduce AI platforms through pilot programmes or funded projects, they struggle to maintain subscriptions after the funding period concludes. This mirrors concerns raised by Patel and Ragolane (2024), who argue that universities often underestimate the hidden costs of AI systems, including subscription renewals, system updates, technical support, and capacity-building initiatives.

In addition to financial concerns, infrastructural challenges, such as limited internet connectivity and lack of access to high-performance devices, were raised, particularly by students from rural areas. This highlights the digital divide that continues to limit the effectiveness of educational innovations. As noted by Haleem et al. (2022), equitable access to both hardware and reliable connectivity is foundational for any technology-driven educational reform. AI tools, which often require substantial computing power and continuous internet access, unintentionally disadvantage students in less-resourced contexts. Supervisors also emphasised the institutional financial burden of maintaining AI platforms beyond the initial rollout phase. While initial access may be funded, ongoing expenses such as licensing, technical support, training, and cybersecurity infrastructure create additional pressures on already stretched university budgets. These financial strains can limit the sustainability of AI integration, leading to a cycle of promising pilot projects that fail to scale. The exclusionary effects of financial barriers were highlighted. One supervisor observed that students unable to afford compatible devices or continuous data costs were effectively excluded from AI-enhanced supervision initiatives. This finding resonates with Laufer et al. (2021), who warn that uncritical adoption of educational technologies without attention to access and equity can reinforce existing inequalities rather than bridge them.

### 3.4 Quality and accuracy of AI output

Another significant challenge associated with the use of AI tools in postgraduate supervision is the questionable quality and accuracy of AI-generated outputs. A senior academic supervisor pointed out that while AI-generated outputs often appear convincing, they can be dangerously inaccurate upon closer inspection: "*AI tools sometimes give results that sound convincing but are completely inaccurate when you double-check the sources. That can mislead a student if they don't verify the information*." A master's candidate student shared a personal experience of receiving poor-quality assistance when attempting to summarise literature using AI tools: "*I used an AI tool to help with summarising literature, but it gave me outdated and irrelevant studies. I had to redo everything manually*." An honours candidate student participant raised a serious concern about fabricated references produced by AI platforms: "*The AI-generated references looked perfect until I tried to find them. Many did not even exist. That is dangerous for any academic work*." Highlighting a different dimension of the issue, an early-career academic supervisor noted the inability of AI tools to engage with local or indigenous knowledge systems: "*These tools are not context-aware. When dealing with local South African data or indigenous knowledge systems, the AI simply does not know what to do with it.*" A senior academic supervisor warned about the broader risk of students' critical thinking being compromised by blind trust in AI-generated content: "*If we are not careful, AI will start shaping how students interpret information, even if the data is flawed. We need to teach critical evaluation more than ever.*" Similarly, a doctoral candidate described confusion resulting from conflicting advice between AI-generated suggestions and their supervisor's guidance: "*I relied on an AI tool to help check my methodology section, but the suggestions it gave contradicted my supervisor's advice. It created confusion instead of clarity.*"

The data underscore a critical and escalating concern regarding the quality and accuracy of AI-generated outputs in postgraduate supervision. Participants articulated how convincingly written yet factually inaccurate outputs, fabricated references, outdated materials, and a lack of contextual sensitivity pose significant risks to academic integrity and quality. Supervisors expressed that while AI-generated outputs often appear polished, they frequently contain factual inaccuracies that could mislead unsuspecting students. This observation aligns with findings by Dwivedi et al. (2023), who caution that AI systems, such as GPT models, often produce plausible nonsense—text that sounds credible but is factually incorrect or misleading. In a research-intensive environment like postgraduate study, such errors can compromise the validity of scholarly work if students fail to critically verify AI-generated content. Students' experiences of receiving outdated or irrelevant summaries from AI tools reflect broader limitations in how AI accesses and updates information. As Kumar et al. (2024) note, many AI models are trained on static datasets, meaning their knowledge is frozen at a certain point, which leads to the risk of citing obsolete research, a significant issue in fast-evolving academic fields. Fabricated references emerged as a particularly serious problem. Students recounted instances where AI tools generated convincing-looking yet non-existent citations, a phenomenon known as AI hallucination. This not only undermines academic credibility but could also lead to severe breaches of academic honesty policies if students unknowingly utilise such fabricated sources in formal submissions.

Supervisors raised important concerns regarding AI's lack of contextual and localised knowledge, particularly concerning indigenous knowledge systems and region-specific research. AI's inherent biases towards dominant Western knowledge paradigms and its inability to meaningfully engage with local contexts have been well-documented. This limitation poses a critical challenge for postgraduate research in diverse contexts such as South Africa, where local data, indigenous methodologies, and culturally sensitive approaches are essential. Another alarming risk is the potential erosion of critical thinking. If students accept AI outputs uncritically, they may internalise flawed information and reproduce it in their work, ultimately weakening their academic rigour. Zhai et al. (2024) emphasise that as AI becomes increasingly embedded in educational practice, there is an urgent need to cultivate critical digital literacy—

the ability to question, verify, and reflect on machine-generated outputs rather than passively accept them. Students' confusion due to conflicting advice from AI tools versus their supervisors underscores a deeper problem: AI is not yet sophisticated enough to replace the nuanced, context-sensitive, dialogical nature of human supervision. As Akgun and Greenhow (2022) argue, genuine education involves a level of emotional, ethical, and contextual intelligence that current AI simply cannot replicate.

### 3.5 Ethical concerns, cultural and contextual sensitivity

Ethical concerns, cultural misalignment, and lack of contextual sensitivity emerged as critical risks when utilising AI tools in postgraduate supervision. One of the senior supervisors expressed concern about the cultural limitations embedded within AI technologies: "*Most of these AI platforms are developed in the West, so they do not always align with our cultural or ethical research standards here in South Africa*." Similarly, a master's candidate voiced anxiety about the potential ethical risks related to confidentiality and sensitive data management: "*I worry that using AI tools could lead to accidental breaches of confidentiality, especially when dealing with sensitive community-based data*." An early-career supervisor highlighted the inability of AI systems to accurately interpret cultural nuances in qualitative research: "*AI does not understand the cultural nuances behind some of our qualitative interviews. It flattens everything and removes the human element*." From the honours candidate's perspective, there was also dissatisfaction with the ethical guidance provided by AI tools, which was often too generic and misaligned with local practices: "*I once asked an AI tool to help with ethical considerations, and it gave generic Western norms that did not apply to our local protocols*." An experienced supervisor cautioned about the risks of students bypassing critical ethical processes due to overreliance on AI assistance: "*There is a risk that students could rely on AI without understanding the ethical approval process, which could lead to serious violations*." A doctoral candidate emphasised the broader issue of AI outputs ignoring local realities and linguistic diversity: "*AI outputs often ignore the local context and language, which is problematic when our research is grounded in South African communities and realities*."

One of the key issues identified by the participants is the cultural limitations of AI technologies. This concern is consistent with research by Ofosu-Asare (2024), which highlighted that AI tools trained on datasets predominantly from Western contexts may be culturally insensitive, reinforcing biases that exclude non-Western perspectives. In the South African context, where research often engages with diverse communities, these limitations are even more pronounced, particularly when AI tools are applied to research involving qualitative methodologies, such as interviews with marginalised or indigenous populations. Furthermore, participants expressed anxiety over ethical risks, particularly related to confidentiality and the handling of sensitive data. This aligns with broader discussions in the literature, where scholars such as Paul (2024) argue that AI systems, particularly those involving cloud-based platforms, pose privacy risks due to potential data breaches or improper handling of sensitive information. In postgraduate research, where data confidentiality is paramount, these concerns are valid and need to be addressed to ensure the ethical integrity of academic work.

Another dimension of the ethical risks discussed is the inability of AI tools to interpret cultural nuances in qualitative research. This limitation highlights the inherent challenges associated with employing AI in fields such as social science and the humanities, where cultural sensitivity is paramount. As De Freitas et al. (2023) suggest, AI tools are not adept at recognising the complexity of human experiences, particularly those rooted in specific cultural or social contexts. In South Africa, where research frequently addresses issues such as inequality, historical trauma, and indigenous knowledge systems, this can result in misinterpretations or oversimplifications of the data.

Concerns regarding the ethical guidance provided by AI tools, which are often too generic and misaligned with local research practices, further underscore the necessity for contextual sensitivity in AI design. This critique aligns with Hagendorff (2022), who argues that AI tools can perpetuate ethical blind spots because they are frequently not designed to account for the ethical standards of varying global contexts. Given that ethical standards can differ significantly across countries and cultures, reliance solely on AI for guidance may lead to unethical practices, particularly in areas involving human subjects or vulnerable communities. Furthermore, the risk of circumventing ethical approval processes due to students' overreliance on AI tools presents another serious concern. Supervisors have warned that students may overlook the necessary steps to obtain ethical clearance for their research if they rely excessively on AI-generated content. This issue has been raised by Osasona et al. (2024), who note that while AI tools can expedite research processes, they should not replace the rigorous, human-driven ethical review processes required for conducting responsible research. Linguistic diversity and local realities emerge as concerns when AI tools fail to account for the local context. This aligns with Zhai et al. (2024), who stress that AI's lack of understanding of diverse linguistic contexts can lead to misinterpretations or the exclusion of non-dominant languages and knowledge systems.

### 3.6 Resistance to change

Resistance to change was highlighted as a major challenge in the adoption of AI tools for postgraduate supervision. A senior lecturer reflected on the scepticism among academic staff regarding AI integration: "*Many of my colleagues are still sceptical about AI; they see it as a threat rather than a support tool*." A master's candidate shared a similar perspective, explaining how some supervisors' feelings of insecurity hinder AI adoption: "*Some supervisors feel that using AI undermines their expertise, so they avoid it completely*." An early-career academic pointed out the broader institutional resistance to innovation: "T*here is a general resistance at the institutional level; they are slow to adopt anything new, especially if it challenges traditional academic practices*." Echoing these sentiments, a doctoral candidate expressed frustration about how fear of change hampers progress: "*It is frustrating because even though AI could help us work smarter, the fear of change holds everyone back*."

Participants articulated that scepticism, fear of undermining traditional expertise, and institutional inertia impede the adoption of AI technologies in academic practices. A supervisor's reflection on the scepticism surrounding AI adoption is consistent with existing literature addressing the barriers to technology integration in educational contexts. As noted by Mirbabaie et al. (2020), academic staff frequently perceive new technologies, such as AI, as a threat to their established practices and professional identities. This fear of disruption is particularly pronounced among those who may feel that AI could challenge their expertise or diminish their role in the research process. The sentiment that AI could "undermine expertise" was echoed by both students and supervisors in the data, underscoring the human element of academic work that many educators consider irreplaceable by technology. In particular, Sobaih et al. (2025) argue that such resistance is exacerbated by a lack of familiarity with AI tools and discomfort regarding their perceived inability to replicate the nuanced judgement and insight that human experts contribute to academic work. The apprehension surrounding AI as a tool that could undermine academic authority also aligns with concerns raised by Singun (2025), who discusses educators' reluctance to embrace digital tools due to perceived threats to their autonomy and academic freedom. In many instances, the introduction of AI tools is viewed as an external imposition on well-established academic processes, thereby generating friction between traditional academic norms and technological innovation. The data presentation indicates that supervisors, particularly those entrenched in traditional pedagogical methods, may entirely avoid using AI tools to safeguard their expertise and maintain control over academic supervision.

Institutional resistance, as highlighted by a supervisor, represents another significant challenge. Higher education institutions often exhibit institutional inertia, characterised by a slow adoption of new technologies that heavily relies on established systems and structures. This resistance is further compounded by the bureaucratic and hierarchical nature of many academic institutions, which can stifle innovation. The level of institutional resistance frequently arises from resource constraints, a lack of training opportunities, and uncertainty regarding the potential benefits of AI. As Pramjeeth and Ramgovind (2024) argue, higher education institutions must transition from a reactive to a proactive stance, adopting policies that actively promote the integration of AI into teaching and research practices to surmount this inertia. The fear of change within academic environments, as expressed by students in the data, reflects a broader reluctance to disrupt the status quo. While AI could present opportunities to work more efficiently, as one student noted, the fear of the unknown and the perceived risks associated with integrating AI into academic workflows can lead to a lack of enthusiasm for its adoption. Kraus et al. (2022) assert that technological change is often met with resistance because it necessitates shifts not only in individual behaviours but also in organisational culture and mindset. In this context, both supervisors and students exhibit hesitance towards embracing AI tools due to the uncertainty surrounding their effectiveness and the potential costs associated with adapting to new technologies.

## 4. Conclusions

This study examined the challenges and risks associated with the utilisation of artificial intelligence (AI) tools for postgraduate supervision within South African higher education contexts. The findings reveal that, while AI tools offer opportunities for enhancing efficiency and accessibility, their integration into postgraduate supervision is fraught with significant challenges. Principal among these are issues of cost and sustainability, inadequate digital literacy and institutional support, overdependence on AI leading to compromised critical thinking, concerns regarding the quality and accuracy of AI-generated outputs, and ethical and cultural insensitivity of AI systems. The research demonstrates that without careful management, AI tools may exacerbate existing inequalities, undermine academic development, and dilute the quality and integrity of postgraduate education. Furthermore, the study highlights that the successful integration of AI tools necessitates more than mere technical adoption; it requires pedagogical redesign, critical engagement, and a commitment to equity, ethics, and contextual relevance. As South African higher education institutions progress towards the incorporation of digital technologies in supervision, a nuanced, critical, and supportive approach will be essential to mitigate risks and foster meaningful educational outcomes.

### 4.1 Recommendations

Universities are urged to assume proactive responsibility for providing subsidised access to AI tools, reliable hardware, and stable internet connectivity for postgraduate students and supervisors, with particular emphasis on individuals from disadvantaged backgrounds. This can be realised through targeted funding schemes, partnerships with technology companies, and internal budget allocations prioritising digital equity. Beyond improving access, the social implications of this recommendation encompass the reduction of digital inequality, the fostering of equitable participation in research, and the prevention of the exacerbation of socio-economic divides within higher education. Practically, sustained institutional investment guarantees that access to AI does not conclude with short-term pilot projects but instead becomes a permanent feature of the research ecosystem. Implementing this recommendation is essential for creating an inclusive research environment where all students can meaningfully leverage AI tools to enhance the quality, efficiency, and innovation of their academic work.

It is recommended that postgraduate schools, in collaboration with digital learning departments and ethics committees, design and deliver formal, ongoing training programmes aimed at developing both the technical competence and critical engagement capacities of supervisors and students regarding the use of AI. These programmes should extend beyond operational training to include modules on ethical standards, verification of AI outputs, responsible referencing, and the maintenance of academic integrity. Clear institutional guidelines and policies should also regulate the appropriate use of AI tools in postgraduate research. The social implications of this recommendation encompass the fostering of a culture of ethical scholarship, the prevention of academic misconduct, and the promotion of responsible digital citizenship within the academic community. Practically, such training equips researchers with the skills to navigate AI-driven environments confidently, reduces inappropriate reliance on AI, and strengthens the supervisory relationship by ensuring a shared understanding of AI's role and limitations.

Universities should integrate critical digital literacy into postgraduate curricula, led by curriculum developers, supervisors, and academic staff, to enable students to critically evaluate, question, and verify AI-generated information. Institutions should also advocate for the development and adoption of culturally sensitive AI systems that incorporate diverse, local, and indigenous knowledge frameworks, potentially through collaborations with AI developers and researchers in the humanities and social sciences. Establishing regular monitoring and evaluation systems, supported by quality assurance units, will assist institutions in tracking AI's impact on supervision, identifying unintended consequences, and refining guidelines over time. The social implications of this recommendation include the promotion of culturally grounded knowledge production, the safeguarding of linguistic and epistemic diversity, and the assurance that AI technologies reflect the lived realities of South African communities. Practically, this approach ensures that AI serves as a tool for empowerment, enhancing research relevance, improving decision-making, and strengthening academic inclusivity, rather than perpetuating bias or homogenising knowledge.

### 4.2 Limitations of study

While the phenomenological, constructivist design produced rich, contextually grounded insights into postgraduate students' and supervisors' experiences with AI tools, several limitations should be acknowledged. The purposive sample of twenty participants (ten students and ten supervisors) from selected South African universities limits the study's generalisability beyond the sampled institutions and disciplines. Additionally, the findings reflect participants' self-reported perceptions rather than observed practices. The reliance on semi-structured interviews introduced the possibility of social desirability and recall biases. Furthermore, the sole use of qualitative content analysis without methodological triangulation (e.g., document analysis or observation) may have constrained the breadth of evidence. Researcher positionality within a constructivist lens could also have influenced data interpretation. Given the rapid evolution of AI technologies, the findings are time-bound and may not fully capture subsequent developments or platform changes. These limitations are discussed in relation to the study design and sample details in the manuscript.

The study suggests further enquiries on:

- How higher education institutions could design contextually responsive policies and support frameworks for the ethical, equitable, and sustainable use of AI tools in postgraduate supervision.
- Postgraduate students' and supervisors' experiences of developing critical AI literacy skills and how these skills impact research quality and academic identity formation in higher education.

## 5. Declarations

**Funding:** This research did not receive any external funding.
**Conflicts of Interest:** The author declares no conflict of interest.
**Use of Artificial Intelligence:** The current work was created without the assistance of artificial intelligence technologies, as confirmed by the author.

## References

Ade-Ibijola, A., & Okonkwo, C. (2023). Artificial intelligence in Africa: Emerging challenges. In *Responsible AI in Africa: Challenges and opportunities* (pp. 101–117). Springer International Publishing.

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K–12 settings. *AI and Ethics, 2*(3), 431–440. https://doi.org/10.1007/s43681-021-00096-7

Al-Zahrani, A. M. (2024). Unveiling the shadows: Beyond the hype of AI in education. *Heliyon, 10*(9), e30696. https://doi.org/10.1016/j.heliyon.2024.e30696

Al-Zahrani, A. M., & Alasmari, T. M. (2024). Exploring the impact of artificial intelligence on higher education: The dynamics of ethical, social, and educational implications. *Humanities and Social Sciences Communications, 11*(1), 1–12. https://doi.org/10.1057/s41599-024-03432-4

Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., ... & Walker, K. (2020). Purposive sampling: Complex or simple? Research case examples. *Journal of Research in Nursing, 25*(8), 652–661. https://doi.org/10.1177/1744987120927206

Clarke, A., Healy, K., Lynch, D., & Featherstone, G. (2023). The use of a constructivist grounded theory method is a good fit for social work research. *International Journal of Qualitative Methods, 22*, 1–10. https://doi.org/10.1177/16094069231186257

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches.* Sage Publications.

De Freitas, J., Agarwal, S., Schmitt, B., & Haslam, N. (2023). Psychological factors underlying attitudes toward AI tools. *Nature Human Behaviour, 7*(11), 1845–1854. https://doi.org/10.1038/s41562-023-01734-2

Delcker, J., Heil, J., Ifenthaler, D., Seufert, S., & Spirgi, L. (2024). First-year students' AI competence as a predictor for intended and de facto use of AI tools for supporting learning processes in higher education. *International Journal of Educational Technology in Higher Education, 21*(1), 1–13. https://doi.org/10.1186/s41239-024-00452-7

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R. (2023). "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges, and implications of generative conversational AI for research, practice, and policy. *International Journal of Information Management, 71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Filiz, O., Kaya, M. H., & Adiguzel, T. (2025). Teachers and AI: Understanding the factors influencing AI integration in K–12 education. *Education and Information Technologies*. Advance online publication. https://doi.org/10.1007/s10639-025-13463-2

Gasaymeh, A. M. M., Beirat, M. A., & Abu Qbeita, A. A. A. (2024). University students' insights into generative artificial intelligence (AI) writing tools. *Education Sciences, 14*(10), 1–18. https://doi.org/10.3390/educsci14101062

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies, 15*(1), 1–28. https://doi.org/10.3390/soc15010006

Gill, M. J. (2020). Phenomenology as a qualitative methodology. In *Qualitative analysis: Eight approaches* (pp. 73–94).

Hagendorff, T. (2022). Blind spots in AI ethics. *AI and Ethics, 2*(4), 851–867. https://doi.org/10.1007/s43681-021-00122-8

Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers, 3*, 275–285. https://doi.org/10.1016/j.susoc.2022.05.004

Kraus, S., Durst, S., Ferreira, J. J., Veiga, P., Kailer, N., & Weinmann, A. (2022). Digital transformation in business and management research: An overview of the current status quo. *International Journal of Information Management, 63*, 102466. https://doi.org/10.1016/j.ijinfomgt.2021.102466

Kumar, V., Ashraf, A. R., & Nadeem, W. (2024). AI-powered marketing: What, where, and how? *International Journal of Information Management, 77*, 102783. https://doi.org/10.1016/j.ijinfomgt.2024.102783

Laufer, M., Leiser, A., Deacon, B., Perrin de Brichambaut, P., Fecher, B., Kobsda, C., & Hesse, F. (2021). Digital higher education: A divider or bridge builder? Leadership perspectives on edtech in a COVID-19 reality. *International Journal of Educational Technology in Higher Education, 18*(1), 1–17. https://doi.org/10.1186/s41239-021-00287-6

Lim, W. M. (2024). What is qualitative research? An overview and guidelines. *Australasian Marketing Journal*. Advance online publication. https://doi.org/10.1177/14413582241264622

Mirbabaie, M., Brünker, F., Möllmann, N. R., & Stieglitz, S. (2022). The rise of artificial intelligence—Understanding the AI identity threat at the workplace. *Electronic Markets, 32*(2), 1–27. https://doi.org/10.1007/s12525-021-00496-x

Ofosu-Asare, Y. (2024). Cognitive imperialism in artificial intelligence: Counteracting bias with indigenous epistemologies. *AI & Society*. Advance online publication. https://doi.org/10.1007/s00146-024-02065-0

Osasona, F., Amoo, O. O., Atadoga, A., Abrahams, T. O., Farayola, O. A., & Ayinla, B. S. (2024). Reviewing the ethical implications of AI in decision-making processes. *International Journal of Management & Entrepreneurship Research, 6*(2), 322–335.

Patel, S., & Ragolane, M. (2024). The implementation of artificial intelligence in South African higher education institutions: Opportunities and challenges. *Technium Education and Humanities, 9*, 51–65.

Paul, J. (2024). *Privacy and data security concerns in AI.* https://www.researchgate.net/publication/385781993_Privacy_and_data_security_concerns_in_AI

Pramjeeth, S., & Ramgovind, P. (2024). Generative artificial intelligence (AI) tools in higher education: A moral compass for the future? *African Journal of Inter/Multidisciplinary Studies, 6*(1), 1–13. https://doi.org/10.51415/ajims.v6i1.1560

Raheem, B. R., Anjum, F., & Ghafar, Z. N. (2023). Exploring the profound impact of artificial intelligence applications (QuillBot, Grammarly, and ChatGPT) on English academic writing: A systematic review. *International Journal of Integrative Research, 1*(10), 599–622. https://doi.org/10.59890/ijir.v1i10.366

Roche, C., Wall, P. J., & Lewis, D. (2023). Ethics and diversity in artificial intelligence policies, strategies, and initiatives. *AI and Ethics, 3*(4), 1095–1115. https://doi.org/10.1007/s43681-022-00218-9

Ruslin, R., Mashuri, S., Rasak, M. S. A., Alhabsyi, F., & Syam, H. (2022). Semi-structured interview: A methodological reflection on the development of a qualitative research instrument in educational studies. *IOSR Journal of Research & Method in Education, 12*(1), 22–29. https://doi.org/10.9790/7388-1201052229

Sabiteka, M., Yu, X., & Sun, C. (2025). Toward sustainable education: A contextualised model for educational technology adoption in developing countries. *Sustainability, 17*(8), 1–44. https://doi.org/10.3390/su17083592

Singun, A. J. (2025). Unveiling the barriers to digital transformation in higher education institutions: A systematic literature review. *Discover Education, 4*(1), 1–41. https://doi.org/10.1007/s44217-025-00430-9

Sobaih, A. E. E., Chaibi, A., Brini, R., & Abdelghani Ibrahim, T. M. (2025). Unlocking patient resistance to AI in healthcare: A psychological exploration. *European Journal of Investigation in Health, Psychology and Education, 15*(1), 1–27. https://doi.org/10.3390/ejihpe15010006

Southworth, J., Migliaccio, K., Glover, J., Glover, J. N., Reed, D., McCarty, C., ... & Thomas, A. (2023). Developing a model for AI across the curriculum: Transforming the higher education landscape via innovation in AI literacy. *Computers and Education: Artificial Intelligence, 4*, 100127. https://doi.org/10.1016/j.caeai.2023.100127

Turin, T. C., Raihan, M., & Chowdhury, N. (2024). Paradigms of approaches to research. *Bangabandhu Sheikh Mujib Medical University Journal, 17*(2), e73973. https://doi.org/10.3329/bsmmuj.v17i2.73973

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*(1), 28. https://doi.org/10.1186/s40561-024-00316-7

Zhang, Y., & Wildemuth, B. M. (2009). Qualitative analysis of content. *Applications of social research methods to questions in information and library science*, *308*(319), 1-12.

Zhang, Y., & Wildemuth, B. M. (2009). Qualitative analysis of content. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (pp. 308–319). Libraries Unlimited.

Zhou, X., Zhang, J., & Chan, C. (2024). Unveiling students' experiences and perceptions of artificial intelligence usage in higher education. *Journal of University Teaching and Learning Practice, 21*(6), 126–145. https://qmro.qmul.ac.uk/xmlui/handle/123456789/96258

CHAPTER EIGHT

# Potential AI-based Use of Fuzzy Cognitive Mapping in Postgraduate Supervision in Higher Education

**Vincent R. Nyirenda**[1]
**Charles B. Chisanga**[2]

AFFILIATIONS

1&2School of Natural Resources, The Copperbelt University, Kitwe, Zambia.

CORRESPONDENCE

Email: vincent.nyirenda@cbu.ac.zm

REFERENCE

Nyirenda, V. R., & Chisanga, C. B. (2026). Potential AI-based use of fuzzy cognitive mapping in postgraduate supervision in higher education. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 130–145). ERRCD Forum. https://doi.org/10.38140/obp4-2026-08



**Abstract:** Supervision in higher education is a complex and evolving process that necessitates adaptive, evidence-based decision-making to effectively guide postgraduate students. Conventional supervisory models often encounter difficulties in addressing uncertainties and the non-linear dynamics inherent in academic mentorship. This chapter examines the AI-based application of Fuzzy Cognitive Mapping (FCM) as an innovative framework to enhance supervisory practices by integrating expert insights, student progress data, and institutional guidelines within a structured yet flexible system. Employing a mental model approach, the study utilises fuzzy logic principles to simulate supervisory scenarios and assess causal relationships among critical factors, such as student motivation, research complexity, institutional support, and mentor–mentee engagement. The FCM-based framework enables supervisors to visualise interdependencies between variables, predict outcomes, and dynamically adjust mentoring strategies. Mixed methods, combining quantitative and qualitative data, are employed. Findings indicate that FCM enhances supervisory efficiency by promoting proactive interventions, improving communication, and supporting continuous monitoring of mentor–mentee relationships. Furthermore, the model advances a data-driven and transparent approach to supervision, minimising subjectivity while preserving contextual flexibility. By operationalising cognitive and computational intelligence, this chapter illustrates how FCM can bridge gaps between qualitative judgement and quantitative assessment in higher education supervision. The study contributes to emerging scholarship on artificial intelligence applications in academic contexts, underscoring the potential of cognitive modelling in improving student outcomes. It concludes by emphasising the necessity of empirical validation and the integration of adaptive mental models into institutional supervisory frameworks to strengthen postgraduate research management and mentoring effectiveness.

***Keywords:*** Artificial intelligence, cognitive mapping, decision-making, higher education, mental models, mentor-mentee relationship.

## 1. Introduction

Postgraduate supervision within higher education is a multi-faceted and dynamic process that requires adaptive decision-making to enhance students' academic experiences, particularly through the utilisation of Artificial Intelligence (AI)-powered tools (Thong et al., 2025). These AI tools are continuously evolving (Michel-Villarreal et al., 2023). AI applications present both opportunities and challenges, particularly concerning graduate research, which must be operationalised through comprehensive training, collaboration, and the establishment of AI policies (Rajab et al., 2025). Furthermore, AI's transformative potential renders it suitable for impactful research evaluations (Arsalan et al., 2025), particularly when utilising predictive models (Hoyos et al., 2023), and serves as an essential resource for monitoring student progress (Zhu & Zhang, 2023). The application of AI in student instruction and assessment demonstrates promising contributions to higher education (Hooda et al., 2022; Liang et al., 2025).

Conventional supervisory models typically operate within rigid frameworks that fail to account for the uncertainties and non-linear nature of academic mentorship in higher education, which are often exacerbated by time constraints (Impola, 2023). Additionally, as higher education diversifies, the changing needs of students, the complexities of research problems, and the evolving institutional academic policies necessitate more flexible and data-driven supervisory approaches. One such promising methodology for achieving this adaptability is Fuzzy Cognitive Mapping (FCM)—a computational technique for modelling, analysing, predicting, and decision-making concerning complex systems through co-production and visualisation (Bakhtavar et al., 2021; Barbrook-Johnson & Penn, 2022). FCMs have been developed to incorporate significant advancements in artificial intelligence, such as machine learning and deep learning, particularly in addressing uncertainty-linked decision-making (Apostolopoulos & Groumpos, 2023). FCMs serve as decision-making support tools through semi-quantitative modelling of complex systems (Barbrook-Johnson & Penn, 2022; Mehryar & Surminski, 2022; Borrero-Domínguez & Escobar-Rodríguez, 2023), informed by collective intelligence (Gray et al., 2020). FCMs have been applied across various domains, including risk analysis (e.g., Bakhtavar et al., 2021), infrastructure construction (e.g., Chen et al., 2024), academic research (e.g., Borrero-Domínguez & Escobar-Rodríguez, 2023), participatory research and decision-making (Sarmiento et al., 2024), nursing research (Andersson & Silver, 2019), and medical sciences (e.g., Apostolopoulos et al., 2024).

The application of computational intelligence, particularly Fuzzy Cognitive Maps (FCMs), in postgraduate supervision within higher education offers a systematic approach to model mentor-mentee relationships, anticipate challenges, identify critical success pathways, and provide timely interventions. This methodology incorporates various components, including expert and collective knowledge, student progress data, and institutional policies or guidelines, to facilitate structured and dynamic real-time decision-making within transformative postgraduate supervision in higher education (Nagaraj et al., 2023). FCMs can also be utilised in conjunction with other methodologies, such as Techniques for Order Preference by Similarity to Ideal Solution (TOPSIS) and Strengths, Weaknesses, Opportunities, and Threats (SWOT) analysis (Baykasoğlu & İlker Gölcük, 2015). Furthermore, FCMs function as a hybrid artificial intelligence tool, integrating expert knowledge, causal reasoning, and fuzzy logic to model complex, dynamic systems while ensuring interpretability and transparency (Apostolopoulos & Groumpos, 2023; Orang et al., 2023). Recent studies illustrate that the integration of FCMs with machine learning, optimisation algorithms, or data-driven techniques enhances predictive accuracy and scenario simulation, thereby providing robust decision support in environmental and socio-ecological contexts (Duhayyim et al., 2023; Shrivastava & Shukla, 2025).

Contemporary research concerning academic mentorship emphasises the significance of mentor-mentee engagement, student motivation, and institutional support in influencing student mental health and learning outcomes (Rasul et al., 2024). However, conventional models present challenges for students, as these frameworks lack a structured mechanism to visualise the complex, dynamic, and interdependent relationships among these components. The utilisation of FCMs in postgraduate supervision enables a proactive approach to co-assess and respond to evolving student needs. This inclusive process is likely to enhance mentorship effectiveness and student performance, thereby supporting the programmatic Theory of Change (ToC) (Reinholz & Andrews, 2020). The ToC posits that the supervision process, accompanied by a series of activities, should culminate in predefined student success goals, such as positive changes in students' learning, mental health, critical thinking, and research skills.

This chapter aims to contribute to postgraduate supervision in higher education through the use of AI-based FCMs. By co-producing an FCM-based framework, this chapter demonstrates beneficial causal relationships and provides supervisors with predictive tools to effectively guide their decision-making processes. It considers multiple situations in postgraduate supervision but does not, by any means, claim comprehensiveness for all the possible and existing situations in higher education. Therefore, the chapter provides illustrations as examples only. Individual supervisors and students need to co-examine their own supervision situations and co-produce the models in various scenarios, together with possible beneficial interventions.

### 1.1 Statement of the problem

Postgraduate supervision is a highly dynamic process that requires a nuanced approach to mentorship. It supports the United Nations' Sustainable Development Goals (SDGs), particularly SDG 4 (Quality Education). Conventional supervision methods often struggle to manage the uncertainties and risks associated with student motivation, research complexity, and institutional support, resulting in inefficiencies and dissatisfaction among students and funders (Deeley et al., 2019; Skinner et al., 2022). In current research on postgraduate supervision, there is an emphasis on the importance of mentor-mentee engagement; however, limited efforts have been made to operationalise these interactions within a structured and predictive framework (Gamage et al., 2021; Goh & Richardson, 2024). Consequently, the inadequate or lack of use of adaptable decision-support tools hampers the supervision process in addressing diverse student needs and academic expectations.

Furthermore, FCMs provide an additional approach to conventional supervision models to tackle these challenges by integrating various components influencing postgraduate supervision into computational models. By incorporating expert-driven weights and fuzzy logic principles (e.g., the use of IF-THEN rules and integrating categorical variables such as low, medium, and high, using a degree of truth that lies between 0 and 1), FCMs enable the simulation of multiple supervisory scenarios, offering supervisors a proactive mechanism to enhance mentorship effectiveness. Therefore, this chapter aims to illustrate the applicability of FCMs in postgraduate supervision, demonstrating their potential to improve decision-making, engagement strategies, and overall student success at the higher education level.

### 1.2 Research question

In this chapter, the following research question is explored to demonstrate the potential use of FCMs in postgraduate supervision in higher education: *How can AI-reinforced FCMs be used to model key variables that shape effective mentor–mentee relationships and supervisory decision-making in postgraduate education?*

## 2. Materials and Methods

As this chapter generically explores the potential use of FCMs in postgraduate supervision within higher education, the data collection is not linked to any specific institution. Instead, it broadly draws on the expert knowledge of authors to provide a discursive framework for supervisors (mentors) and students (mentees). Furthermore, it is envisaged that users will be inspired to co-develop their own relevant mental models, co-evaluate them, and further refine their applicability in different scenarios while conducting their postgraduate supervision.

The data used in this chapter was collated from multiple sources, including online supervision protocols (e.g., https://www.timeshighereducation.com) and long-term interactions with staff and students at The Copperbelt University over a decade, since 2014, during the authors' course of duty. This period (2014-2025) coincided with the AI boom in several sectors, including higher education (Pisica et al., 2023; Mah & Groß, 2024; An et al., 2025). The data collected from these sources included the identification of key components, the direction of effects between components (whether positive or negative), the magnitude of the effects between components, and participants' perspectives regarding postgraduate supervision. The variables used in the study were generated and consolidated with the help of a generative AI model, GPT-4o, in October 2024. Following the compilation of the variables, expert knowledge was sought from 27 participants (15 academic staff and 12 graduate students) whose opinions aligned with those of the extant variables. In the next step, involving data collection in November 2024, 36 participants (16 academic staff and 20 graduate students from different disciplines) were randomly selected and asked to independently establish relationships between the variables and weight them against a 5-point Likert scale, represented by 1-very weak, 2-somewhat weak, 3-neutral, 4-somewhat strong, and 5-very strong relationship.

The study was ethically overseen by The Copperbelt University, in accordance with the 2024 Helsinki declaration on human participants in research. Mixed methods, combining quantitative and qualitative data, were employed as a case study at The Copperbelt University. Only academics with experience supervising graduate students and students in their final research year participated in the exercise. The average weight of the magnitude of each relationship between components or elements was then integrated into the FCM.

### 2.1 Model elicitation

The shared perceptions of the participating individuals formed the key areas of agreement within the shared mental models. Differences in the data were not assessed, as they are beyond the scope of this chapter. Intentionally, no questionnaires or guides were employed, as it was intended to build a framework that would be illustrative only. The key elements or components for model development were identified from the sources until the saturation point, where components became repetitive. A 1-5 Likert scale was used to assign weights by gauging participants' collective assertions and providing the magnitude. Clear directional dimensions (negative or positive), annotated as – or +, respectively, between pairs of elements in relationships, were also provided by the participants. The ratings were transformed into a -1 to +1 number scale to meet the model's requirements.

### 2.2 Data analysis

The data were analysed using Mental Modeller (https://www.mentalmodeler.com/) – an open FCM software for shared mental models that includes scenario analysis. Mental Modeller has four key features: components with relationships that include magnitudes and dimensions; adjacency matrices; model validation; and scenario analysis. The scenario analysis helps to highlight anticipated system performance when dealing with different parameterised settings. Model validation encompasses the network's structural properties (indegrees, outdegrees, and centrality), where indegree reflects the number of incoming connections, outdegree represents outgoing connections, and centrality indicates the importance of a component within the system.

Corrective measures in supervision performance are implemented based on decision-making informed by the preferred scenarios, which are determined through mentor-mentee engagement. As system influencers, the drivers identified in the validation interface, when increased or decreased, lead to predictable scenario outcomes in the supervision processes. Receivers, however, tend to be influenced or impacted by other elements, while ordinaries are regular, unspecialised elements that create domino effects within the system. Manipulating these types of elements will help to comprehend the potential outcomes of interventions in real time. The number of scenarios that can be created is as many as desired in a particular setting, and some examples will be provided. By refining alterations in system elements through stakeholders' in-depth consultations, scenario analysis can yield comparably higher value outcomes (Goswami et al., 2021).

### 2.3 The scope

The illustrations of FCM presented in this chapter are based on online resources, expert knowledge, and stakeholder engagement spanning over a decade. They adopt broad-brush approaches in identifying components or elements for illustrative purposes only. The FCM connects to an array of focus, inter alia: (i) student progress monitoring through real-time tracking from multiple perspectives, such as the levels of acquisition of research skills; (ii) adaptive decision-making support for personalised student supervision; (iii) prediction of student outcomes regarding potential delays, drop-out risks, and success probabilities; (iv) improving supervisor-student communication by reducing misunderstandings and enhancing research collaboration; (v) optimising resource allocation, such as mentorship time and research funding; and (vi) identifying and addressing stress- and wellbeing-related effects through appropriate social, emotional, and psychological remedial mechanisms.

# 3. Results

## 3.1 Model validation

There are specific measures to note. In the exemplar shown in Figure 1, there were 23 components (elements), 27 connections, a density of 0.053, 1.173 connections per component, 12 driver components, 5 receiver components, 6 ordinary components, and a complexity score of 0.41. These various model measures may differ depending on the co-created situation relevant to postgraduate supervision.

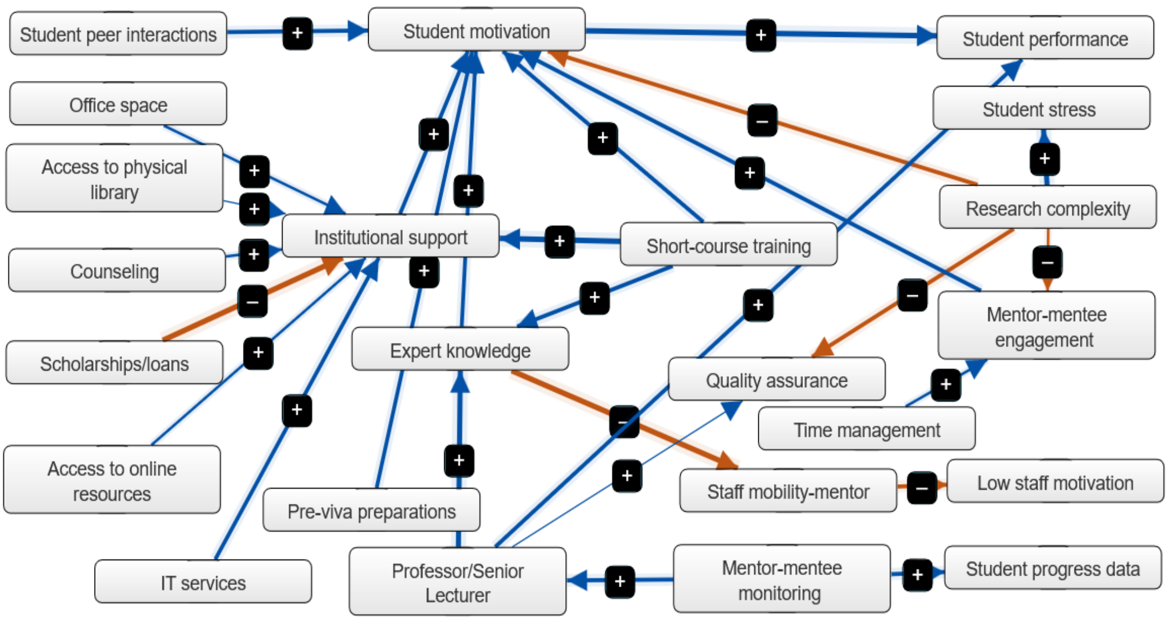



***Figure 1:*** *Illustration of AI-retrieved components and their FCM-generated relationships*

## 3.2 Component performance

Table 1 reflects the features presented in the model (Figure 1) and provides specific measures of the network. In this network, the software indicates that student motivation has the highest sum of indegree.

*Table 1: Adjacency matrix for indegrees, outdegrees, and centrality of individual components.*

| Components | Indegree | Outdegree | Centrality | Type |
|---|---|---|---|---|
| Student motivation | 5.18 | 0.88 | 6.06 | Ordinary |
| Institutional support | 3.75 | 0.68 | 4.43 | Ordinary |
| Expert knowledge | 1.71 | 1.67 | 3.38 | Ordinary |
| Research complexity | 0 | 2.61 | 2.61 | Driver |
| Short-course training | 0 | 2.43 | 2.43 | Driver |
| Professors or senior lecturers | 0.53 | 1.74 | 2.27 | Ordinary |
| Student performance | 1.62 | 0 | 1.62 | Receiver |
| Staff mobility-mentor | 0.84 | 0.6 | 1.44 | Ordinary |
| Mentor-mentee engagement | 0.65 | 0.66 | 1.31 | Ordinary |
| Mentor-mentee monitoring | 0 | 1.23 | 1.23 | Driver |
| Student stress | 1.00 | 0 | 1.00 | Receiver |
| Scholarship or loans | 0 | 0.99 | 0.99 | Driver |
| Pre-viva preparations | 0 | 0.81 | 0.81 | Driver |
| Student peer interactions | 0 | 0.80 | 0.80 | Driver |
| Quality assurance | 0.80 | 0 | 0.80 | Receiver |
| Student progress data | 0.70 | 0 | 0.70 | Receiver |
| IT services | 0 | 0.65 | 0.65 | Driver |
| Low staff motivation | 0.60 | 0 | 0.60 | Receiver |
| Access to online resources | 0 | 0.47 | 0.47 | Driver |
| Time management | 0 | 0.44 | 0.44 | Driver |
| Counselling | 0 | 0.37 | 0.37 | Driver |
| Office space | 0 | 0.22 | 0.22 | Driver |
| Access to physical library | 0 | 0.13 | 0.13 | Driver |

Outdegree and centrality values serve as a mediating component. Other important mediating components include the availability of institutional support, expert knowledge, and supervisors (professors or senior lecturers). The key drivers of postgraduate supervision, however, are research complexity, short-course training, mentor-mentee monitoring, the availability of scholarships or education loans, pre-viva preparations, and student peer interactions. The potential outcomes (receivers) encompass student performance, levels of student stress (mental health), quality assurance, and student progress data. Together, all these elements function to support the levels of postgraduate supervision.

### 3.3 Scenario development

For illustration purposes, a 5-component scenario is implemented in Mental Modeller to develop a model with the outcomes shown in Figure 2. The five components used in the model include student motivation, research complexity, mentor-mentee engagement, availability of student scholarships or education loans, and professors or senior lecturers. Consequently, the model suggests that postgraduate supervision is likely to improve through student performance, expert knowledge, institutional support and quality assurance, and reduced student stress (mental health), by increasing student motivation, mentor-mentee engagement, availability of student scholarships or education loans, and professors or senior lecturers, while decreasing research complexity.

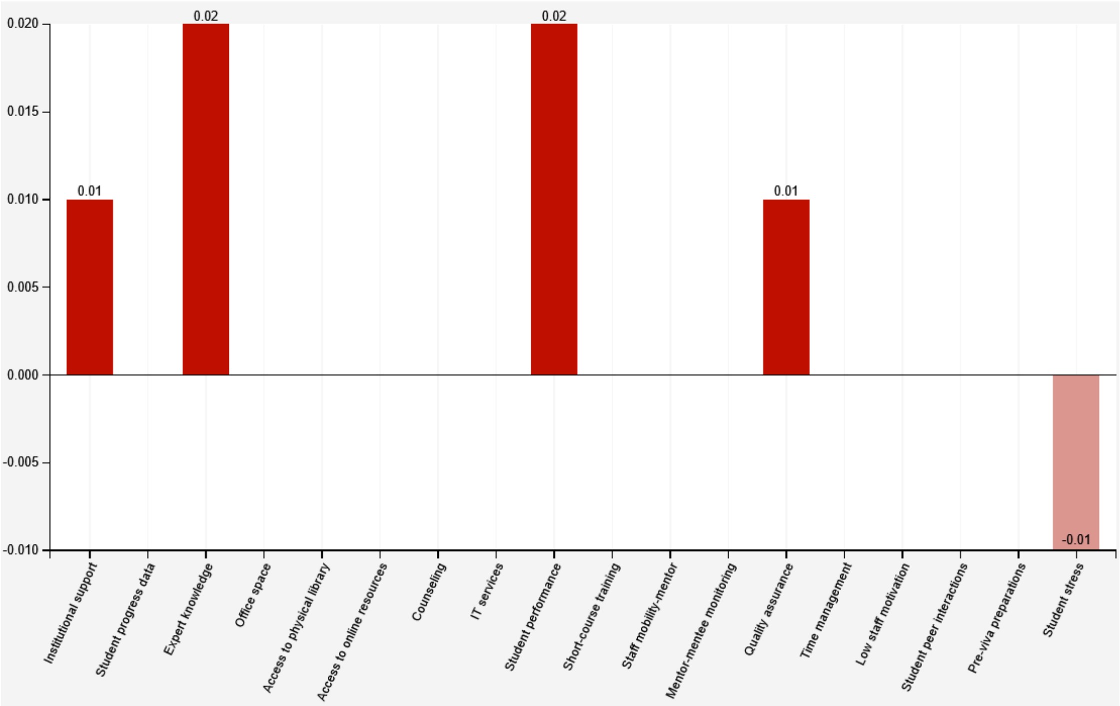



***Figure 2:*** *An example of parameterising*

Figure 2 is an example of parameterising (increasing or reducing) values of particular components to obtain the desired results in scenario planning. In this case, student motivation, mentor-mentee engagement, scholarships or loans, and the availability of professors or senior lecturers are increased, while research complexity is reduced based on the principle of parsimony.

## 4. Discussion

In our model illustrations, student motivation is the most influential mediating component in postgraduate supervision. This outcome is consistent with past research, emphasising its role in academic persistence and research productivity among postgraduate students (Litalien & Guay, 2015; Shen & Jiang, 2021; Khosa et al., 2023). In addition, the model identifies other key mediators, such as institutional support, expert knowledge, and supervisor involvement, reinforcing previous studies that underscore these components as essential to effective supervision, particularly when addressing uncertainties in supervision processes (Albertyn & Bennett, 2020). Notably, the study delineates critical drivers of postgraduate supervision, including research complexity, short-course training, mentor-mentee monitoring, availability of scholarships or education loans, pre-viva preparations, and student peer interactions. These drivers align with the literature advocating for structured coaching or training programmes and financial support, which are vital for the well-being and resilience of postgraduate students, thus enhancing research outcomes (Casey et al., 2022).

The application of AI-based FCMs using Mental Modeler offers an approach to understanding postgraduate supervision dynamics amid the uncertainties faced by students in their studies. The five-component scenario—comprising student motivation, research complexity, mentor-mentee engagement, scholarships or education loans, and supervisory expertise—yields outcomes that corroborate various aspects of supervisory practice (Taylor & Kiley, 2024). The exploration of these aspects should be conducted inductively and comprehensively, especially with the use of AI, and personalised for individual students throughout the study period. Furthermore, our illustrative model suggests that reducing research complexity positively impacts supervision quality, complementing studies that advocate for structured, well-scaffolded research processes to mitigate cognitive overload (Kiley, 2015; Auhl & Bain, 2023). FCM can be applied to define challenges and articulate solutions in postgraduate supervision (Andersson & Silver, 2019). It can also assist in formalising stakeholder knowledge, supporting learning in supervision, and promoting remedial actions.

FCMs facilitate a nuanced representation of interdependencies among key components or elements for supervisory decision-making. The ability to visualise causal relationships, with a sense of direction and weighted magnitude of effects, provides supervisors with a data-driven basis for real-time interventions. This supports computational pedagogical modelling for improved supervisory quality (Yao & Lin, 2023; Bond et al., 2024). Furthermore, scenario analysis in Mental Modeler for FCMs helps predict supervisory outcomes, particularly those fostering supervisory effectiveness through more targeted dialogue and decision-making. This approach permits supervisors to tailor mentorship programmes and allocate resources effectively. This aligns with recent advancements advocating for AI and machine learning in academic decision-making within higher education (Hu et al., 2024; Kalnina et al., 2024). Andersson and Silver (2019) elaborated that FCMs can be used to formalise steps that connect and evaluate students' progress in research and studies in achieving their learning objectives.

However, despite their merits, the application of FCMs in postgraduate supervision may present certain challenges, namely: (i) the complexity of capturing qualitative supervisory interactions in a computational model remains a limitation; (ii) although the average outcomes from FCMs are reliable (Aminpour et al., 2021), there is a reliance on accurate expert-driven weight assignments for model calibration; and (iii) additionally, the dynamism of supervision and variations in supervisory styles across disciplines and cultural contexts may limit the generalisability of FCM-based frameworks (Sarmiento et al., 2024). These constraints highlight the need for individualised applications and further refinement and empirical validation of FCM applications in postgraduate supervision.

Interestingly, this study diverges from some prior literature in its assertion that research complexity should be actively reduced to enhance supervision effectiveness, especially when students find it extremely challenging to handle. However, conventional perspectives postulate that exposure to complex research environments fosters intellectual resilience (Wisker &

Robinson, 2016). The discrepancy may stem from differing conceptualisations of research complexity—where structured complexity benefits students, while unstructured, excessive difficulty may hinder progress. Therefore, future studies could further delineate these distinctions, refining the role of research complexity in postgraduate supervision models.

## 5. Conclusions

This chapter elucidates that fuzzy cognitive maps (FCMs) serve as a potent instrument for revealing the fundamental dynamics that govern postgraduate supervision. The model identifies student motivation as the most significant mediating component, bolstered by institutional support, expert knowledge, and supervisory involvement. It additionally uncovers that research complexity, financial support, mentor-mentee monitoring, and training function as critical drivers that influence outcomes such as student performance, stress levels, quality assurance, and research progress. Moreover, the findings indicate that the quality of supervision can be enhanced by augmenting motivational and supervisory enablers, while concurrently managing excessive research complexity, thus emphasising a distinction between productive and obstructive forms of complexity.

The utilisation of FCMs via Mental Modeller exemplifies how AI-supported computational modelling can enhance traditional supervisory practices by visualising causal structures, quantifying influence pathways, and facilitating predictive scenario analysis. This analytical capability furnishes supervisors with a data-driven foundation for timely, targeted interventions and more personalised mentorship strategies. By formalising stakeholder knowledge and integrating AI-aligned decision-support tools, institutions can improve supervisory consistency, bolster student resilience, and strengthen overall research outcomes.

### 5.1 Social and practical implications

The findings of this study convey significant social implications by highlighting the central role of student motivation, institutional support, and manageable research complexity in fostering healthier and more equitable postgraduate research environments. By recognising how these factors influence student stress, performance, and persistence, higher education institutions can better design support systems that alleviate mental health burdens, promote inclusive academic participation, and enhance students' overall well-being. Enhancing motivational supports and reducing unnecessary research barriers can also contribute to improved retention and completion rates, thereby advancing broader societal goals related to human capital development and knowledge production.

Practically, the integration of FCMs and AI-enabled modelling provides supervisors and higher education institutions with a structured, evidence-based tool for improving decision-making in real time. The ability to visualise causal pathways and test supervisory scenarios enables more targeted mentorship, efficient allocation of resources, and earlier identification of students who may be at risk of delayed progress or disengagement. This approach supports more transparent supervision practices, strengthens quality assurance processes, and offers a scalable method for training supervisors. Therefore, higher education institutions adopting FCMs can enhance supervisory effectiveness, streamline postgraduate management, and embed data-driven strategies into routine academic practice.

### 5.2 Recommendations

To enhance postgraduate supervision, several targeted recommendations emerge from this study. First, student motivation should be bolstered through structured support programmes and mentorship initiatives designed to sustain engagement throughout the research process. Research complexity should be simplified by providing clear guidelines, phased research targets, and well-scaffolded methodologies to minimise cognitive overload. Mentor-mentee engagement must be strengthened through regular feedback sessions and monitoring mechanisms that ensure close supervision and timely interventions. Institutions of higher education should also expand postgraduate support services to improve access to funding, training opportunities, and administrative assistance. Increasing financial aid, through scholarships, student loans, research grants, and stipends, remains critical for alleviating the financial pressures that often impede student progress.

Additionally, institutions of higher education should implement short-course training in research methodology and academic writing to equip students with essential competencies, while standardising pre-viva preparation programmes to ensure students are adequately prepared for thesis defence. The inductive adoption of AI-based supervision models, such as FCMs, as part of supervisory practice that embeds evidence-based, AI-informed decision-making in postgraduate research management, can further enhance personalised progress tracking and intervention strategies. Integrating scenario-based planning using tools such as Mental Modeller will help supervisors anticipate challenges and optimise supervision strategies. Developing adaptive, discipline-sensitive supervision frameworks is also essential to accommodate variations in research complexity and cultural contexts. Finally, continuous empirical validation is necessary to refine and strengthen FCM-based supervision models to ensure their long-term relevance and applicability across academic disciplines.

## 6. Declarations

**Funding**: This research did not receive any external funding.
**Acknowledgements**: We are grateful for the contributions of various anonymous individuals in the higher education sector.
**Conflicts of Interest**: The author(s) declare no conflict of interest.
**Use of Artificial Intelligence:** The current work was created with the assistance of artificial intelligence technologies (ChatGPT version 4o) to assist with generating initial potential themes associated with postgraduate supervision.

## References

Albertyn, R., & Bennett, K. (2020). Containing and harnessing uncertainty during postgraduate research supervision. *Higher Education Research and Development*, *40*(4), 661–675. https://doi.org/10.1080/07294360.2020.1775559

Aminpour, P., Gray, S.A., Jetter, A.J., Introne, J.E., Singer, A., & Arlinghaus, R. (2020). Wisdom of stakeholder crowds in complex social–ecological systems. *Nature Sustainability*, *3*, 191–199. https://doi.org/10.1038/s41893-019-0467-z

An, Y., Yu, J.H., & James, S. (2025). Investigating higher education institutions' guidelines and policies regarding the use of generative AI in teaching, learning, research, and administration. *International Journal of Educational Technology in Higher Education*, *22*, 10. https://doi.org/10.1186/s41239-025-00507-3

Apostolopoulos, I. D., & Groumpos, P. P. (2023). Fuzzy cognitive maps: Their role in explainable artificial intelligence. *Applied Sciences*, *13*(6), 3412. https://doi.org/10.3390/app13063412

Apostolopoulos, I., Papandrianos, N.I., Papathanasiou, N.D., & Papageorgiou, E.I. (2024). Fuzzy cognitive map applications in medicine over the last two decades: A review study. *Bioengineering*, *11*(2), 139. https://doi.org/10.3390/bioengineering11020139

Andersson, N., & Silver, H. (2019). Fuzzy cognitive mapping: An old tool with new uses in nursing research. *Journal of Advanced Nursing*, *75*(12), 3823–3830. https://doi.org/10.1111/jan.14192

Arsalan, M. H., Mubin, O., Al Mahmud, A., Khan, I. A., & Hassan, A. J. (2025). Mapping data-driven research impact science: The role of machine learning and artificial intelligence. *Metrics*, *2*(2), 5. https://doi.org/10.3390/metrics2020005

Auhl, G., & Bain, A. (2023). Preparing practitioners for inclusive practice: The challenge of building schema to reduce cognitive load. In C. Boyle & K. A. Allen (Eds.), *Research for inclusive quality education*. Sustainable Development Goals Series. Springer. https://doi.org/10.1007/978-981-16-5908-9_5

Bakhtavar, E., Valipour, M., Yousefi, S., Sadiq, R., & Hewage, K. (2021). Fuzzy cognitive maps in systems risk analysis: A comprehensive review. *Complex and Intelligent Systems*, *7*, 621–637. https://doi.org/10.1007/s40747-020-00228-2

Barbrook-Johnson, P., & Penn, A. S. (2022). *Systems mapping: How to build and use causal models of systems*. Palgrave Macmillan. https://doi.org/10.1007/978-3-031-01919-7_6

Baykasoğlu, A., & Gölcük, İ (2015). Development of a novel multiple-attribute decision making model via fuzzy cognitive maps and hierarchical fuzzy TOPSIS. *Information Sciences*, *301*, 75–98. https://doi.org/10.1016/j.ins.2014.12.048

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S.W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of*

*Educational Technology in Higher Education*, *21*, 4. https://doi.org/10.1186/s41239-023-00436-z

Borrero-Domínguez, C., & Escobar-Rodríguez, T. (2023). Decision support systems in crowdfunding: A fuzzy cognitive maps (FCM) approach. *Decision Support Systems*, *173*, 114000. https://doi.org/10.1016/j.dss.2023.114000

Casey, C., Harvey, O., Taylor, J., Knight, F., & Trenoweth, S. (2022). Exploring the wellbeing and resilience of postgraduate researchers. *Journal of Further and Higher Education*, *46*(6), 850–867. https://doi.org/10.1080/0309877X.2021.2018413

Chen, H., Cheng, S., Qin, Y., Xu, W., & Liu, Y. (2024). Sustainability evaluation of urban large-scale infrastructure construction based on dynamic fuzzy cognitive map. *Journal of Cleaner Production*, *449*, 141774. https://doi.org/10.1016/j.jclepro.2024.141774

Deeley, S.J., Fischbacher-Smith, M., Karadzhov, D., & Koristashevskaya, E. (2019). Exploring the 'wicked' problem of student dissatisfaction with assessment and feedback in higher education. *Higher Education Pedagogies*, *4*(1), 385–405. https://doi.org/10.1080/23752696.2019.1644659

Duhayyim, M., Mohamed, H. G., Alzahrani, J. S., Alabdan, R., Mousa, M., Zamani, A. S., Yaseen, I., & Alsaid, M. I. (2023). Modeling of fuzzy cognitive maps with a metaheuristics-based rainfall prediction system. *Sustainability*, *15*(1), 25. https://doi.org/10.3390/su15010025

Gamage, K. A. A., Perera, D. A. S., & Wijewardena, M. A. D. N. (2021). Mentoring and coaching as a learning technique in higher education: The impact of learning context on student engagement in online learning. *Education Sciences*, *11*(10), 574. https://doi.org/10.3390/educsci11100574

Goh, E., & Richardson, S. (2024). Developing effective mentoring programs in hospitality higher education: A practical perspective using the mentoring framework. *Journal of Teaching in Travel and Tourism*, *24*(2), 179–188. https://doi.org/10.1080/15313220.2024.2304748

Goswami, R., Roy, K., Dutta, S., Ray, K., Sarkar, S., Brahmachari, K., Nanda, M.K., Mainuddin, M., Banerjee, H., Timsina, J., & Majumdar, K. (2021). Multi-faceted impact and outcome of COVID-19 on smallholder agricultural systems: Integrating qualitative research and fuzzy cognitive mapping to explore resilient strategies. *Agricultural Systems,189*, 103051. https://doi.org/10.1016/j.agsy.2021.103051

Gray, S., Aminpour, P., Reza, C., Scyphers, S., Grabowski, J., Murphy Jr., R., Singer, A., Baltaxe, D., Jordan, R., Jetter, A., & Introne, J. (2020). Harnessing the collective intelligence of stakeholders for conservation. *Frontiers in Ecology and Environment*, *18*(8), 465–472. https://doi.org/10.1002/fee.2232

Hooda, M., Rana, C., Dahiya, O., Rizwan, A., & Hossain, M. S. (2022). Artificial intelligence for assessment and feedback to enhance student success in higher education. *Mathematical Problems in Engineering*, *1*, 5215722. https://doi.org/10.1155/2022/5215722

Hoyos, W., Aguilar, J., & Toro, M. (2023). PRV-FCM: An extension of fuzzy cognitive maps for predictive modelling. *Expert Systems with Applications*, *231*, 120729. https://doi.org/10.1016/j.eswa.2023.120729

Hu, S., Ke, F., Vyortkina, D., Hu, P., Luby, S., & O'Shea, J. (2025). Artificial intelligence in higher education: Applications, challenges, and policy development and further considerations. In L.W. Perna (Ed.), *Higher education: Handbook of theory and research* (Vol. 40). Springer. https://doi.org/10.1007/978-3-031-51930-7_13-1

Kalniņa, D., Nīmante, D., & Baranova, S. (2024). Artificial intelligence for higher education: Benefits and challenges for pre-service teachers. *Frontiers in Education*, *9*, 1501819. https://doi.org/10.3389/feduc.2024.1501819

Khosa, A., Wilkin, C., & Burch, S. (2023). PhD students' relatedness, motivation, and well-being with multiple supervisors. *Accounting Education*, *33*(2), 131–163. https://doi.org/10.1080/09639284.2023.2179889

Kiley, M. (2015). 'I didn't have a clue what they were talking about': PhD candidates and theory. *Innovations in Education and Teaching International*, *52*(1), 52–63. https://doi.org/10.1080/14703297.2014.981835

Liang, J., Stephens, J.M., & Brown, G.T.L. (2025). A systematic review of the early impact of artificial intelligence on higher education curriculum, instruction, and assessment. *Frontiers in Education*, *10*, 1522841. https://doi.org/10.3389/feduc.2025.1522841

Litalien, D., & Guay, F. (2015). Dropout intentions in PhD studies: A comprehensive model based on interpersonal relationships and motivational resources. *Contemporary Educational Psychology*, *41*, 218–231. https://doi.org/10.1016/j.cedpsych.2015.03.004

Impola, J.T. (2023). Reconsidering Newtonian temporality in the context of time pressures of higher education. *Studies in Philosophy and Education*, *42*, 431–448. https://doi.org/10.1007/s11217-023-09879-3

Mah, D.K., & Groß, N. (2024). Artificial intelligence in higher education: Exploring faculty use, self-efficacy, distinct profiles, and professional development needs. *International Journal of Educational Technology in Higher Education*, *21*, 58. https://doi.org/10.1186/s41239-024-00490-1

Mehryar, S., & Surminski, S. (2022). Investigating flood resilience perceptions and supporting collective decision-making through fuzzy cognitive mapping. *Science of the Total Environment*, *837*, 155854. https://doi.org/10.1016/j.scitotenv.2022.155854

Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D.E., Thierry-Aguilera, R., & Gerardou, F.S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, *13*(9), 856. https://doi.org/10.3390/educsci13090856

Nagaraj, B.K., Kalaivani, A., Suraj Begum, R., Akila, S., Sachdev, H.K., & Senthil Kumar, N. (2023). The emerging role of artificial intelligence in STEM higher education: A critical review. *International Research Journal of Multidisciplinary Technovation*, *5*(5), 1–19. https://doi.org/10.54392/irjmt2351

Orang, O., Silva, P.C. de Lima e., & Guimarães, F.G. (2023). Time series forecasting using fuzzy cognitive

maps: A survey. *Artificial Intelligence Review*, *56*, 7733–7794. https://doi.org/10.1007/s10462-022-10319-w

Pisica, A. I., Edu, T., Zaharia, R. M., & Zaharia, R. (2023). Implementing artificial intelligence in higher education: Pros and cons from the perspectives of academics. *Societies*, *13*(5), 118. https://doi.org/10.3390/soc13050118

Rasul, T., Nair, S., Kalendra, D., Balaji, M.S., de Oliveira Santini, F., Ladeira, W.J., Rather, R.A., Yasin, N., Rodriguez, R.V., Kokkalis, P., Murad, M.W., & Hossain, M.U. (2024). Enhancing academic integrity among students in GenAI era: A holistic framework. *International Journal of Management Education*, *22*(3), 101041. https://doi.org/10.1016/j.ijme.2024.101041

Rajab, S., Aramadhan, C., & Asad, K.W. (2025). Artificial intelligence adaptation in higher education: Supervisors' efficiency in detecting AI-generated research in postgraduate research. *International Journal of Social Science and Human Research*, *8*(9), 7181–7191. https://doi.org/10.47191/ijsshr/v8-i9-59

Reinholz, D.L., & Andrews, T.C. (2020). Change theory and theory of change: What is the difference anyway? *International Journal of STEM Education*, *7*, 2. https://doi.org/10.1186/s40594-020-0202-3

Sarmiento, I., Cockcroft, A., Dion, A., Belaid, L., Silver, H., Pizarro, K., Pimentel, J., Tratt, E., Skerritt, L., Ghadirian, M.Z., Gagnon-Dufresne, M.-C., & Andersson, N. (2024). Fuzzy cognitive mapping in participatory research and decision making: A practice review. *Archives of Public Health,82*, 76. https://doi.org/10.1186/s13690-024-01303-7

Shen, W., & Jiang, J. (2021). Institutional prestige, academic supervision, and research productivity of international PhD students: Evidence from Chinese returnees. *Journal of Sociology,56*(2). https://doi.org/10.1177/14407833211055225

Shrivastava, V., & Shukla, S. (2025). Illuminating themes of ecovillages by leveraging participatory modelling: A convergent gen-AI and fuzzy cognitive maps approach. *Journal of Environmental Management, 385*, 125650. https://doi.org/10.1016/j.jenvman.2025.125650

Skinner, E.A., Kindermann, T.A., Vollet, J.W., & Rickert, N.P. (2022). Complex social ecologies and the development of academic motivation. *Educational Psychology Review,34*, 2129–2165. https://doi.org/10.1007/s10648-022-09714-0

Taylor, S., & Kiley, M. (2024). *A handbook for doctoral supervisors* (3rd ed.). Routledge. https://doi.org/10.4324/9781003432159

Thong, C.L., Atallah, Z., Islam, S., Lim, W.L., & Cherukuri, A.K. (2025). AI-powered tool for doctoral supervision in higher education: A systematic review. *Journal of Information and Knowledge Management,24*(2), 1–27. https://doi.org/10.1142/S0219649225300013

Yao, D., & Lin, J. (2023). Identifying key factors influencing teaching quality: A computational pedagogy approach. *Systems, 11*(9), 455. https://doi.org/10.3390/systems11090455

Wisker, G., & Robinson, G. (2016). Supervisor well-being and identity: Challenges and strategies. *International Journal for Researcher Development, 7*(2), 123–140. https://doi.org/10.1108/IJRD-03-2016-0006

Zhu, Z., & Zhang, L. (2023). Artificial intelligence empowers postgraduate education ecologically sustainable development model construction. *Sustainability, 15*(7), 6157. https://doi.org/10.3390/su15076157

CHAPTER NINE

# Fostering Intellectual Autonomy in AI-Mediated Postgraduate Supervision: A Gamified Approach

**Tirivanhu Muchuweni**[1]
**Zingiswa Jojo**[2]

AFFILIATIONS

[1&2]Department of Secondary and Post-School Education, Faculty of Education, Rhodes University, Makhanda, South Africa.

CORRESPONDENCE

Email: Tirivanhu.Muchuweni@gmail.com

REFERENCE

Muchuweni, T., & Jojo, Z. (2026). Fostering intellectual autonomy in AI-mediated postgraduate supervision: A gamified approach. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 146–165). ERRCD Forum. https://doi.org/10.38140/obp4-2026-09



**Abstract:** This chapter examines how gamification strategies facilitate self-directed learning, intellectual autonomy, and academic growth within AI-mediated postgraduate supervision. Employing a systematic literature review methodology in accordance with PRISMA 2020 standards, the chapter synthesises research that investigates the integration of gamification and artificial intelligence in postgraduate education. Key strategies, including milestone-based progression, collaborative problem-solving, and interactive feedback, are analysed as mechanisms for fostering autonomy and intrinsic motivation among postgraduate students. The review indicates that structured challenges and rewards enhance engagement, reduce over-reliance on AI tools, and support the development of complex research skills. Furthermore, gamification serves as a catalyst for professional growth, with collaborative and reflective practices enhancing critical thinking and resilience. Ethical considerations, such as algorithmic bias, fairness, and the necessity of maintaining human-centred mentorship, are also examined. The chapter concludes that the integration of gamification into AI-mediated supervision frameworks promotes creativity, independence, and intellectual rigour, providing valuable insights for educators and institutions seeking to align technological innovation with the enduring values of postgraduate mentorship.

***Keywords:*** AI in education, gamification, intellectual autonomy, postgraduate supervision, self-directed learning.

## 1. Introduction

The integration of artificial intelligence (AI) into postgraduate supervision is fundamentally reshaping academic mentorship by introducing more personalised and efficient support systems (Dai et al., 2023; Makokotlela, 2024; Sim et al., 2023). Tools such as ChatGPT and EndNote streamline research tasks, support academic writing and referencing, and enhance overall productivity (Luckin, 2024; Zawacki-Richter et al., 2019; Chauke et al., 2024; Cui, 2024; Kızıltaş, 2025). Nevertheless, these advancements also raise concerns regarding the erosion of intellectual autonomy, which remains a core pillar of academic and professional development (Veletsianos & Kimmons, 2012; Zawacki-Richter et al., 2019; Deci & Ryan, 1985; Deci & Ryan, 2000). Scholars (Marengo & Pange, 2024; Ferris, 2011; Bond et al., 2020) caution that excessive reliance on AI may diminish students' opportunities to cultivate critical thinking, independent learning, and reflective skills. Furthermore, the evolving relationship between supervisors and students necessitates that supervisors strike a balance between the rapid advancement of technology and the provision of human guidance that fosters deep learning (Luckin, 2024; Laurillard, 2013;

Albertyn & Bennett, 2021; Emilsson & Johnsson, 2007; Oparinde, 2021). Some studies (Dai et al., 2023; Cowling et al., 2023; Chauke et al., 2024; Kızıltaş, 2025) suggest that while students value AI support, a significant risk exists that they may begin to rely on such technologies even for tasks that necessitate original thought and clear understanding.

As AI becomes increasingly commonplace in postgraduate education, there is a pressing need to reconsider how students learn, receive feedback, and engage with their supervisors (Qudsi, 2024; Rivera, 2021; Cui, 2024; Dai et al., 2023; Sim et al., 2023). Most AI systems are designed to prioritise clarity and efficiency, which may enhance superficial academic performance but can impede deeper cognitive engagement (Selwyn, 2019; Chauke et al., 2024; Kızıltaş, 2025). In the context of postgraduate research, where sustained reflection and conceptual development are critical, this limitation is particularly concerning (Luckin, 2024; Bell, 2016; Jones, 2017; Willcoxson, 1998). Selwyn (2019) and Zawacki-Richter et al. (2019) warn that AI-generated feedback may undermine the dialogic and relational aspects of supervision (Albertyn & Bennett, 2021; Daramola, 2021; Maistry, 2015). Research conducted by Dai et al. (2023) and Selwyn (2019) indicates that students tend to produce more reflective and meaningful work when AI support is augmented with feedback from their supervisors, rather than relying exclusively on AI (Cowling et al., 2023; Sim et al., 2023).

Hamari, Koivisto, and Sarsa (2014), Deterding, Dixon, Khaled, and Nacke (2011), Subhash and Cudney (2018), and Muchuweni, Jojo, and Kariyana (2025) demonstrate that gamification can enhance motivation and support self-directed learning. In postgraduate contexts, where learners manage complex, multi-phase research tasks, gamified structures can provide essential scaffolding and maintain momentum (Rivera, 2021; Dicheva et al., 2015; Indriasari et al., 2020; Landers, 2014; Ortiz-Rojas et al., 2025; Zainuddin et al., 2024). Features such as progress tracking, achievement systems, and leaderboards assist students in monitoring their development, encouraging accountability, and recognising milestones (Müller & Mildenberger, 2021; Looyestyn et al., 2017; Landers, 2014; Li et al., 2024; Ortiz-Rojas et al., 2025; Venter & de Wet, 2024). Integrated feedback mechanisms within gamification facilitate iterative improvement and continuous reflection (Qudsi, 2024; Black & Wiliam, 2009; Olsher et al., 2016; Ramadhan et al., 2024). Findings from a recent systematic literature review further substantiate these patterns, with evidence demonstrating that Quizizz enhances student engagement, strengthens motivation, improves academic performance, and supports formative assessment (Muchuweni et al., 2025, p. 119). Several studies (Hamari et al., 2014; Dicheva et al., 2015; Subhash & Cudney, 2018; Omodan & Marongwe, 2024) indicate that students respond positively when gamification is employed to scaffold meaningful checkpoints, such as proposal submissions or reflective journaling.

Poorly designed gamified systems may lead students to focus on rewards rather than meaningful learning (Nicholson, 2015; Li et al., 2024; Seaborn & Fels, 2015). When these strategies lack intrinsic alignment, they can promote shallow engagement and discourage deeper inquiry (Deci

& Ryan, 1985; Leon et al., 2015; Ng et al., 2012; Reeve, 2012; Standage, 2023). However, the convergence of artificial intelligence (AI) and gamification in postgraduate supervision also introduces complex ethical considerations (Floridi & Cowls, 2022; Buckley, Doyle, & Doyle, 2017; Cui, 2024; Dai et al., 2023; Xu, 2025). AI systems may perpetuate algorithmic bias by favouring dominant academic norms or linguistic styles, which can disadvantage students from non-traditional backgrounds (Winkler & Söllner, 2018; Floridi & Cowls, 2022; Bond et al., 2020; Zawacki-Richter et al., 2019). Furthermore, unequal access to digital tools and platform variability across institutions may exacerbate existing educational inequities (Zawacki-Richter et al., 2019; Cullen et al., 2020; Hidayat & Firmanti, 2024; Viberg et al., 2020). Gamification structures that overemphasise competition or fixed metrics can diminish inclusivity and weaken intrinsic motivation (Buckley et al., 2017; Li et al., 2024; Sailer & Homner, 2020; Yığ & Sezgin, 2021). Ethically grounded supervision should remain adaptive, inclusive, and attentive to the unintended consequences of digital tools (Albertyn & Bennett, 2021; Maistry, 2015; Oparinde, 2021).

Despite the increasing body of research on artificial intelligence (AI) and gamification as discrete domains, there remains a dearth of studies that explore their combined impact on postgraduate supervision (Bond et al., 2020; Dichev & Dicheva, 2017; Majuri et al., 2018; Rahmi et al., 2025; Triantafyllou et al., 2025). The majority of gamification research concentrates on undergraduate settings, primarily with short-term engagement objectives (Subhash & Cudney, 2018; Dicheva et al., 2015; Indriasari et al., 2020; Lathwesen & Belova, 2021; Ortiz-Rojas et al., 2025; Sánchez-Arévalo et al., 2025; Yllana-Prieto et al., 2025), while studies addressing AI in education frequently prioritise administrative or technical functionalities over pedagogical depth (Dai et al., 2023; Selwyn, 2019; Chauke et al., 2024; Cowling et al., 2023; Cui, 2024; Kızıltaş, 2025; Sim et al., 2023). This separation has contributed to a fragmented understanding of the interactions between these tools and their potential to support or impede postgraduate development (Bond et al., 2020; Li et al., 2024; Rahmi et al., 2025; Triantafyllou et al., 2025). Furthermore, many ethical and pedagogical implications of these technologies remain under-theorised (Marengo & Pange, 2024; Floridi & Cowls, 2022; Looyestyn et al., 2017; Seaborn & Fels, 2015; Xu, 2025).

While concerns such as algorithmic bias, cognitive overload, and gamification fatigue are acknowledged, they are seldom examined in relation to sustained academic growth (Buckley et al., 2017; Rivera, 2021; Dichev & Dicheva, 2017; Hamari et al., 2014; Li et al., 2024; Sailer & Homner, 2020; Triantafyllou et al., 2025). Much of the literature continues to emphasise task performance rather than the long-term development of critical thinking, resilience, and research identity (Veletsianos & Kimmons, 2012; Bond et al., 2020; Landers, 2014; Leon et al., 2015; Ng et al., 2012; Ortiz-Rojas et al., 2025; Rahmi et al., 2025; Sánchez-Arévalo et al., 2025). This gap between the utilisation of technology and pedagogical objectives highlights the necessity for more integrative research that considers student development, equity, and sustained learning (Bond et al., 2020; Li et al., 2024; Triantafyllou et al., 2025). Institutions may adopt AI and gamification primarily for operational efficiency, often overlooking their deeper academic ramifications (Dockendorff & Gómez Zaccarelli, 2025; Oulaich, 2019; Shumba, 2024; Viberg et al., 2020). In the absence of careful planning, these technologies risk valuing convenience over mentorship, reflection, and intellectual profundity (Dichev & Dicheva, 2017; Hamari et al., 2014; Landers, 2014). Consequently, this chapter employs the PRISMA 2020 framework (Page et al., 2021) to synthesise contemporary literature and provide guidance for the design of learner-centred, ethical, and autonomy-supportive models of AI-mediated postgraduate supervision.

## 1.1 Problem statement

While AI-supported postgraduate supervision enables more personalised feedback and efficient administration (Aoun, 2017), it also raises concerns about students' independence and decision-making in research (Selwyn, 2019; Melisa et al., 2025). Gamification has been proposed as a complementary strategy to address these risks by enhancing motivation, structuring progression, and promoting self-directed learning in AI-mediated environments (Dicheva et al., 2015; Hamari et al., 2014). Its role in supporting intellectual independence at the postgraduate level is still not well studied (Subhash & Cudney, 2018). This study addresses a critical gap in evidence on how gamification can be strategically applied to counterbalance the cognitive and ethical risks associated with over-reliance on AI in postgraduate supervision. If this gap remains unaddressed, students may become more dependent on AI, their intellectual autonomy may weaken, and institutions may overlook important ethical risks in supervision.

### *1.1.1 Research question*

This study investigates the role of gamification in fostering intellectual autonomy within AI-mediated postgraduate supervision and is structured around the following guiding question:

- How does the integration of gamification into AI-mediated postgraduate supervision influence students' intellectual autonomy, including their critical thinking and independent research skills?

## 1.2 Theoretical framework: Self-Determination theory

Self-Determination Theory (SDT), developed by Deci and Ryan (1985), is utilised in this chapter to explain how AI tools and gamified structures influence postgraduate supervision. SDT posits that learning improves when three basic needs are supported: autonomy, competence, and relatedness (Deci & Ryan, 1985; Ryan & Deci, 2017). Autonomy involves making one's own decisions in learning (Deci & Ryan, 1985; Ryan & Deci, 2017). While AI tools can assist with routine tasks, students may lose autonomy if they rely too heavily on automated suggestions or overly rigid digital systems (Dai et al., 2023; Selwyn, 2019; Chauke et al., 2024; Albertyn & Bennett, 2021; Oparinde, 2021). Competence refers to the feeling of being capable and witnessing improvement (Deci & Ryan, 1985; Ryan & Deci, 2017). Gamified elements, such as progress tracking and milestone badges, can enhance competence by providing clear steps and visible progress (Subhash & Cudney, 2018; Dicheva et al., 2015). Research also indicates that meaningful feedback helps students recognise their growth and build confidence (Reeve, 2023; Standage, 2023). Relatedness pertains to feeling supported by others (Deci & Ryan, 1985; Ryan & Deci, 2017). Human feedback encourages deeper thinking and reflection, something that AI cannot fully replicate (Laurillard, 2013; Selwyn, 2019; Emilsson & Johnsson, 2007; Maistry, 2015). Studies demonstrate that students develop more effectively when AI support is combined with supervisor feedback (Dai et al., 2023; Laurillard, 2013). Overall, SDT helps explain how AI and gamification can either support or undermine autonomy, competence, and relatedness in postgraduate learning. Figure 1 illustrates how these three needs relate to AI tools and gamified supervision.


***Figure 1:*** *SDT components and their connection to AI and gamified postgraduate supervision*

## 2. Materials and Methods

This chapter draws on a structured review of academic literature, employing the PRISMA 2020 approach (Page et al., 2021), to examine how gamified strategies can encourage independent thinking within AI-supported postgraduate supervision. The review synthesised empirical and theoretical research published between 2014 and 2025 on the integration of gamification and AI technologies in higher education, with a focus on postgraduate contexts. A structured search was conducted across three academic databases (ERIC, ScienceDirect, and SpringerLink). Search terms were created to capture studies linking gamification, AI-mediated supervision, and intellectual autonomy. Examples included "gamification AND higher education," "gamification AND AI-mediated supervision," "intellectual autonomy AND postgraduate education," and "artificial intelligence AND postgraduate supervision." Only peer-reviewed studies written in English and situated in higher education were included. Both empirical and theoretical studies were eligible. The reference lists of key papers were manually checked to identify additional sources, and duplicate studies were removed during the screening process. Figure 2 presents the main search topics and keywords used in this review.

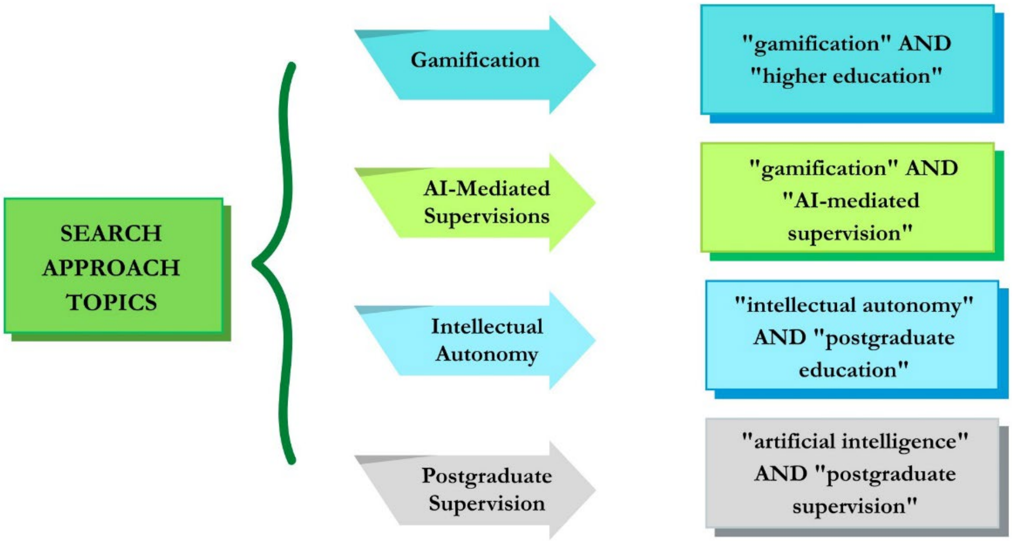



***Figure 2:*** *Search topics and keyword combinations for the review*

Clear inclusion and exclusion criteria guided the selection of studies. Eligible studies examined gamification in higher education, AI-supported supervision, or concepts related to motivation, engagement, or intellectual autonomy. Studies were excluded if they focused solely on undergraduate populations without relevant insights for postgraduate contexts, were not written in English, or did not pertain to higher education. These criteria are summarised in Figure 3.

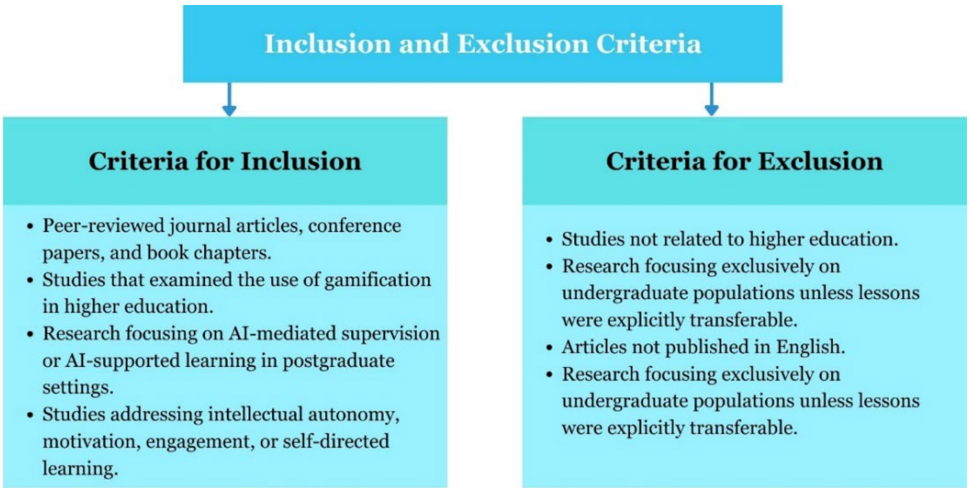



***Figure 3:*** *Summary of inclusion and exclusion criteria used in this review*

The selection process followed the PRISMA 2020 protocol to ensure rigour and transparency. After duplicates were removed, titles and abstracts were screened for relevance based on the predefined criteria, and full-text articles were reviewed to confirm eligibility. A total of 41 studies met the inclusion criteria and were included in the qualitative synthesis. Data extraction was then carried out for all included studies using a structured extraction form. Information recorded for each study included the author and year of publication, country of study, research question, study design and methodology, gamification strategies used, details of AI integration, and findings related to engagement, motivation, intellectual autonomy, and postgraduate supervision. The extracted data were analysed using a narrative synthesis approach to identify common themes and patterns. Figure 4 presents the PRISMA flow diagram that summarises the identification, screening, eligibility, and inclusion stages.

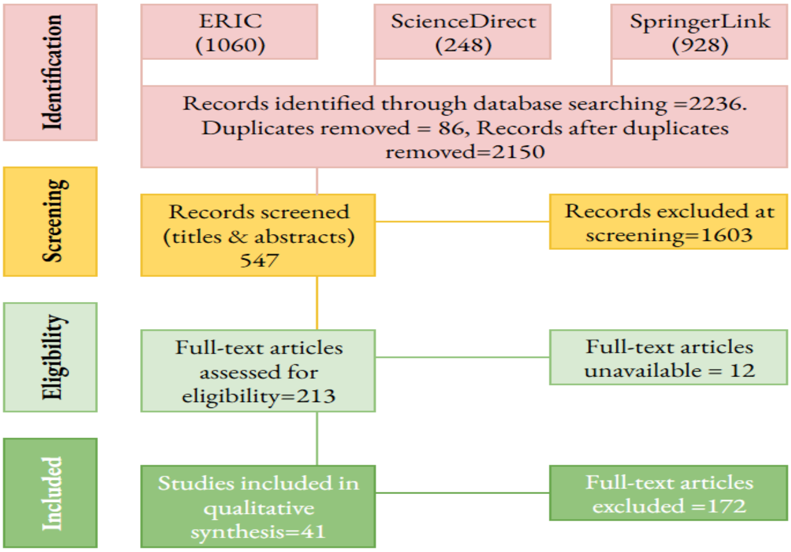



***Figure 4:*** *PRISMA 2020 diagram*

## 3. Presentation of Results

Across the 41 included studies, a few common patterns emerged. Many of the studies indicated that gamification increased student engagement, particularly when combined with AI tools that support feedback or supervision. Elements such as progress tracking, badges, leaderboards, and real-time feedback enhanced motivation and improved interaction with supervision platforms (Hamari et al., 2014; Dicheva et al., 2015). These features helped students stay actively involved, complete research tasks on time, and monitor their progress. Milestone tracking facilitated long-term project management by breaking research into smaller, achievable goals. Leaderboards encouraged peer accountability and friendly competition, while instant feedback helped maintain momentum and reinforce progress. Collectively, these gamification elements contributed to higher levels of task engagement and platform participation. Gamified feedback loops enhance students' engagement with formative feedback in AI-mediated supervision environments (Shute, 2008; Luckin, 2024). Immediate feedback, paired with revision incentives, encourages students to view research as an evolving process rather than a series of isolated tasks.

Well-designed gamified systems motivate students to revise, reflect, and improve their work through iterative learning cycles. These systems align with self-regulated learning principles by reinforcing progress and encouraging repeated engagement with challenging content (Subhash & Cudney, 2018). Evidence from recent gamification research also indicates that students benefit when progress is divided into clear steps and when improvements are recognised over time (Muchuweni et al., 2025). The structure of these feedback loops is illustrated in Figure 5, which outlines the typical cycle of feedback, reflection, and revision in AI-mediated gamified supervision models. However, the effectiveness of feedback depends on its quality and cognitive load. Studies have noted that when feedback is excessive or poorly structured, students may experience cognitive overload or confusion (Carless, 2006), which can result in surface-level revisions or disengagement. To support sustainable learning, gamified supervision should employ tiered feedback. AI systems can provide quick insights early on, while human supervisors offer deeper feedback at key milestones. This layered approach maintains student motivation, manages workload, and promotes meaningful revision.

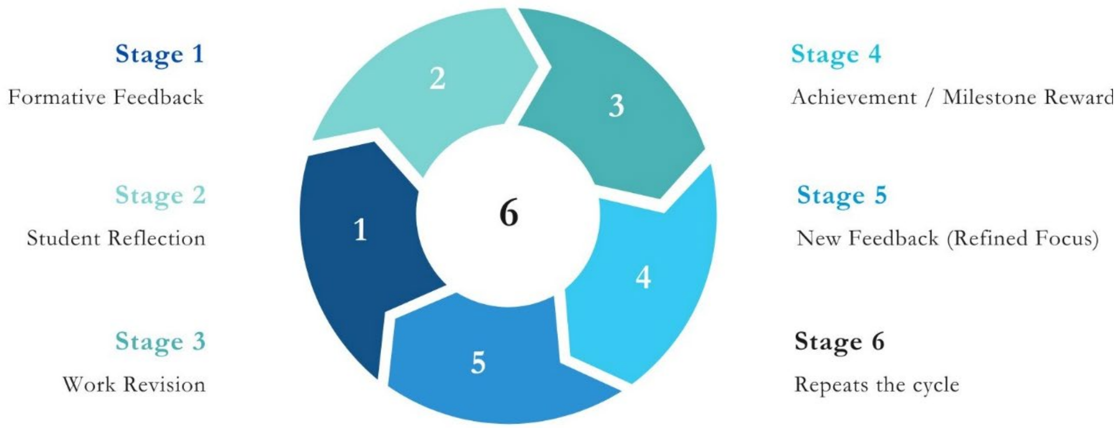



***Figure 5:*** *Gamified Feedback and Learning Iterations in AI-Mediated Postgraduate Supervision*

The reviewed literature consistently emphasised that thoughtfully designed gamified progression systems contribute to improved research milestone completion rates in postgraduate supervision settings. Structured elements such as goal-setting mechanisms, visual progress tracking, and achievement recognition were strongly associated with enhanced student accountability and timely submission of key research components, including literature reviews, data analyses, and thesis drafts (Nicholson, 2015; Omodan & Marongwe, 2024). By breaking complex research projects into smaller, achievable goals, gamified systems promote a sense of continuous advancement. This incremental structure supports better time management, reduces procrastination, and helps sustain motivation over extended research timelines. Progress visualisation tools, such as completion charts and milestone badges, provide immediate reinforcement, encouraging consistent student engagement. However, some studies cautioned that an overemphasis on external rewards could lead students to prioritise task speed over deep academic engagement (Nicholson, 2015; Deci & Ryan, 1985; Buckley et al., 2017). To address this risk, researchers recommended integrating reflective components alongside milestone achievements to maintain academic rigour and critical inquiry. Table 1 summarises the key gamification features identified in the reviewed studies that support research milestone completion.

Ethical and institutional challenges were prominent across the reviewed studies, with recurring concerns related to algorithmic bias, over-reliance on AI feedback, gamification fatigue, cognitive overload, and inequitable access to digital tools (Floridi & Cowls, 2022; Zawacki-Richter et al., 2019). Algorithmic bias was identified as a significant risk, with AI systems sometimes privileging certain research approaches or student profiles based on narrow training datasets. Over-reliance on AI-generated feedback raised concerns about weakening students' critical thinking and independent research skills.

**Table 1:** *Gamification features supporting research milestone completion*

| Gamification Element | Contribution to Research Milestone Completion |
|---|---|
| Structured Goal-Setting | Organises complex research projects into manageable phases |
| Visual Progress Tracking | Provides immediate reinforcement through visible advancement |
| Achievement Badges | Recognises intermediate accomplishments to sustain persistence |
| Reflective Milestone Reviews | Ensures depth of engagement and prevents superficial task focus |

Gamification fatigue, arising from static leaderboards and repetitive rewards, has been linked to reduced motivation, while poorly structured feedback systems risk contributing to cognitive overload. Furthermore, unequal access to AI technologies could exacerbate existing educational inequalities. To address these risks, studies recommend using hybrid supervision models that combine AI support with human mentorship (Luckin, 2024; Laurillard, 2013), improving the diversity of AI training datasets to reduce bias (Floridi & Cowls, 2022), prioritising clear and manageable feedback (Carless, 2006), designing adaptive gamification systems that support autonomy rather than relying solely on rewards (Nicholson, 2015), and ensuring equitable access to digital tools and platforms across institutions (Zawacki-Richter et al., 2019).

In postgraduate supervision, gamification functions very differently from its use in classroom settings. Instead of short quizzes or competitive games, studies such as those by Rivera (2021), Müller and Mildenberger (2021), and Omodan and Marongwe (2024) showed that gamification is employed to structure long-term research tasks, support independent decision-making, and guide students through extended milestones such as proposal development, data collection, and thesis drafting. Nicholson (2015) also emphasised that meaningful gamification supports autonomy by helping students manage progress without relying solely on rewards. These studies demonstrated that gamified elements act as organisational tools rather than classroom activities, as they assist students in monitoring progress, managing timelines, and sustaining motivation over prolonged periods of supervision. This distinction is important because postgraduate learning requires sustained autonomy and deep reflection, rather than rapid classroom engagement. A concise summary of these key ethical challenges and proposed mitigation strategies is provided in Table 2.

***Table 2:*** *Key ethical challenges and recommended mitigations in AI-mediated gamified supervision*

| Ethical Challenge | Description | Recommended Mitigation |
|---|---|---|
| Algorithmic Bias | AI systems may favour certain research approaches or student profiles if trained on narrow datasets. | Regular audits of AI feedback systems; diversify training datasets; include human oversight. |
| Over-Reliance on AI | Excessive dependence on automated feedback may weaken students' critical thinking and research independence. | Maintain hybrid supervision models combining AI feedback with human mentorship. |
| Gamification Fatigue | Prolonged exposure to static leaderboards or reward systems may cause anxiety, disengagement, and reduced motivation. | Implement adaptive gamification strategies that evolve with student progression and offer personalised pathways. |
| Cognitive Overload | Overwhelming students with frequent or poorly prioritised feedback can impair learning efficiency. | Prioritise feedback relevance; stagger notifications; embed structured reflection opportunities. |
| Inequitable Access | Disparities in access to AI tools or digital infrastructure may exacerbate educational inequalities. | Ensure institutional support for equitable access to AI resources and digital tools. |

## 4. Discussion of Findings

The findings show that gamification can support engagement, motivation, and intellectual autonomy in AI-mediated postgraduate supervision, and these patterns can be better understood through Self-Determination Theory (Deci & Ryan, 1985). Across the reviewed studies, gamification is shown to increase student engagement in AI-supported supervision by employing elements such as milestone tracking, leaderboards, badges, and real-time feedback (Hamari et al., 2014; Dicheva et al., 2015). These features encourage more consistent interaction with supervision platforms and help students stay on track with long-term research tasks (Hamari et al., 2014; Dicheva et al., 2015; Muchuweni et al., 2025). Hamari et al. (2014) find that milestone-based tracking fosters a sense of ownership, while Dicheva et al. (2015) emphasise that competitive elements, such as leaderboards, encourage peer interaction and accountability.

Students also demonstrate stronger engagement when they can visualise progress and receive immediate feedback. These findings align with SDT's competence principle, which states that motivation increases when learners can see evidence of their progress. However, several studies caution that motivation may decline over time if gamified systems remain static. This supports the need for adaptive designs that evolve as students advance through their research. The review also indicates that gamification contributes to intellectual autonomy when systems support self-regulation, independent planning, and critical reflection. Structured milestones, personalised pacing, and iterative feedback encourage students to take ownership of their work, reflecting the autonomy component of SDT (Subhash & Cudney, 2018; Shute, 2008). These elements are especially important in postgraduate supervision, where learners must plan months of research, track their progress, and make independent decisions. However, studies warn that designs focused too heavily on points or badges can shift attention away from deep learning. Nicholson (2015) and Deci and Ryan (1985) argue that such designs may lead students to prioritise task completion rather than meaningful academic engagement. In AI-supported postgraduate supervision, gamification therefore works best when it encourages reflection and decision-making rather than simply rewarding speed or task completion. Table 3 summarises the specific gamification elements identified in the reviewed studies that support intellectual autonomy in AI-mediated supervision contexts.

***Table 3:*** *Gamification elements supporting intellectual autonomy among postgraduate students*

| Gamification Strategy | How it Supports Autonomy | Supporting Authors |
|---|---|---|
| Milestone Tracking | Encourages self-paced goal setting and monitoring | Subhash & Cudney (2018) |
| Achievement Badges | Reinforces mastery experiences and self-confidence | Ryan & Deci (2017) |
| Interactive Feedback Loops | Promotes iterative reflection and independent improvement | Shute (2008) |
| Self-Assessment Challenges | Supports critical self-evaluation and autonomy | Vygotsky (1978); Wingate (2006) |

Gamified feedback systems significantly enhance iterative learning processes by promoting active student engagement with formative feedback. A range of studies indicates that immediate and structured feedback increases revision cycles and facilitates deeper critical thinking (Shute, 2008; Luckin, 2024; Muchuweni et al., 2025). Such patterns are consistent with the competence principle of Self-Determination Theory (SDT), as they assist students in developing confidence in revising and improving their work. Incentives such as badges for multiple revision cycles or rewards for enhanced drafts serve to maintain engagement over extended periods of supervision. However, an excess of or poorly organised feedback may introduce cognitive overload, thereby diminishing learning efficiency (Carless, 2006). To mitigate this risk, tiered feedback models are recommended. These systems gradually transition students from frequent AI-generated suggestions to more in-depth feedback from supervisors at critical points, thereby supporting both cognitive manageability and sustained development.

The literature also demonstrates that gamification influences the timely completion of research milestones within postgraduate supervision. Tools such as milestone maps, completion trackers, and achievement systems enhance accountability and assist students in adhering to structured timelines (Nicholson, 2015; Omodan and Marongwe, 2024). By deconstructing large research tasks into manageable steps, these tools facilitate improved time management and reduce procrastination. This approach is aligned with the autonomy principle of SDT, as students are afforded the opportunity to select how they navigate tasks while still perceiving clear progress.

However, an over-reliance on external rewards can distract from meaningful engagement (Deci and Ryan, 1985). Institutions may counter this by integrating reflective components such as research journals or self-assessments, which aid students in connecting milestone completion to deeper academic growth. Ethical considerations also emerge within the literature, as scholars caution that algorithmic bias may manifest in AI feedback systems when training data fails to adequately represent diverse academic backgrounds, potentially disadvantaging certain students (Floridi and Cowls, 2022). Excessive dependence on automated tools may impair critical thinking skills if students rely too heavily on AI-generated suggestions (Zawacki-Richter et al., 2019). Furthermore, gamification fatigue is reported when systems employ repetitive reward structures, resulting in diminished intrinsic motivation over time (Buckley et al., 2017). To address these challenges, researchers advocate for regular audits of AI systems, dynamic reward designs that adapt to students' needs, and sustained human oversight to maintain the relational aspects of supervision.

The findings demonstrate that gamification can significantly enhance student engagement, motivation, and intellectual autonomy when applied judiciously within AI-mediated postgraduate supervision. Interactive feedback systems and structured progression pathways align effectively with Self-Determination Theory (SDT) by fostering autonomy, competence, and relatedness. However, institutions must meticulously manage ethical risks such as bias, over-reliance on automation, and diminished motivation resulting from inadequately designed gamification systems. Supervision models that integrate adaptive gamification, reflective learning tools, ethical oversight, and human mentorship can ensure that AI-driven supervision reinforces postgraduate learning rather than undermining it. The design of such balanced systems is critical for safeguarding the long-term integrity and sustainability of postgraduate research training.

While the findings offer valuable insights, several limitations should be acknowledged. Firstly, publication bias may exist, as studies yielding positive results are more likely to be disseminated. Secondly, the review encompassed only English-language sources, potentially excluding pertinent works from other regions. Thirdly, variations in the definitions and applications of gamification and AI across studies complicated direct comparisons. Notwithstanding these limitations, the review provides significant insights into the influence of gamification and AI on postgraduate supervision and identifies areas meriting further research.

## 5. Conclusions and Recommendations

This chapter consolidates key insights derived from the systematic review and provides recommendations for the integration of gamification into AI-mediated postgraduate supervision. The studies reviewed indicate that gamification has the potential to enhance student engagement, motivation, and intellectual autonomy by offering structured, interactive, and rewarding supervision environments. Features such as milestone tracking, progress visualisation, badges, and timely feedback assist postgraduate students in remaining organised, maintaining motivation, and completing research tasks more consistently. These advantages are particularly pertinent in long-term postgraduate work, where sustained engagement and self-regulation are crucial for academic success. Concurrently, the literature highlights several risks that necessitate careful consideration. An over-reliance on AI-generated feedback may lead to superficial learning, diminished critical thinking, or excessive dependence on automated tools. Algorithmic bias may create inequitable learning environments, especially when AI systems are predicated on narrow or non-representative datasets. Furthermore, gamification fatigue may occur when systems employ repetitive rewards or static design features, consequently undermining intrinsic motivation over time. These concerns underscore the necessity for robust human oversight, ethical safeguards, and ongoing evaluations of AI technologies to uphold academic integrity and foster equitable supervision experiences. Notwithstanding these challenges, the overarching evidence suggests that gamification possesses significant potential to enhance postgraduate research training when integrated within a thoughtfully designed, human-centred supervision model. Effective designs should support intellectual autonomy, encourage iterative learning, and provide opportunities for reflection and self-assessment. Institutions must strike a balance between innovation and responsibility, ensuring that AI tools complement rather than supplant meaningful academic mentorship.

Several recommendations emerge from this review. Institutions are encouraged to implement adaptive gamification designs that enable students to personalise milestones, badges, and progress tracking as they advance through their research. The combination of AI-generated feedback with human mentorship remains essential for preserving depth, reflection, and contextualised guidance. Regular audits of AI systems are necessary to identify and mitigate bias, while clear ethical guidelines should reinforce the irreplaceable role of human judgment in supervision. Embedding reflective practices, such as research journals and self-assessment tools, can further enhance students' intellectual autonomy and self-regulated learning. Future research should investigate the long-term effects of gamified supervision, particularly how dynamic AI-driven gamification systems can adapt to learner progression. Studies are also required to examine how cultural and institutional contexts influence the efficacy of gamified supervision and to explore emerging ethical issues, such as academic integrity and the impact of automated feedback on student decision-making. Assessing faculty readiness and institutional capacity will also be critical for facilitating responsible and sustainable implementation.

## 6. Declarations

**Funding**: This research received no external funding. All findings were developed independently by the authors without financial support from any organisation.

**Acknowledgements:** The authors sincerely thank their families for their support and understanding throughout this work. They also express gratitude to their colleagues for their guidance and professional assistance during the research and writing process. Additionally, the authors acknowledge their academic peers for their constructive feedback and scholarly contributions.

**Conflicts of Interest:** The authors report no conflicts of interest. All interpretations were made independently and ethically, without external influence.

**Use of Artificial Intelligence:** The authors developed all intellectual content, analysis, and arguments independently. Artificial intelligence tools (ChatGPT 5.1) were used only for language refinement and clarity, not for developing the conceptual framework, analysis, or conclusions.

## References

Albertyn, R., & Bennett, K. (2021). Containing and harnessing uncertainty during postgraduate research supervision. *Higher Education Research & Development, 40*(6), 1122–1135. https://doi.org/10.1080/07294360.2020.1775559

Aoun, J. E. (2017). *Robot-proof: Higher education in the age of artificial intelligence.* MIT Press. https://doi.org/10.7551/mitpress/11456.001.0001

Bell, A. (2016). Intercultural postgraduate supervision: Reimagining time, place and knowledge. *Higher Education Research & Development, 35*(2), 227–240. https://doi.org/10.1080/07294360.2016.1106931

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5–31. https://doi.org/10.1007/s11092-008-9068-5

Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *International Journal of Educational Technology in Higher Education, 17*(1), 1–30. https://doi.org/10.1186/s41239-019-0176-8

Buckley, P., Doyle, E., & Doyle, S. (2017). Game on! Students' perceptions of gamified learning. *Educational Technology & Society, 20*(3), 1–10.

Carless, D. (2006). Differing perceptions in the feedback process. *Studies in Higher Education, 31*(2), 219–233. https://doi.org/10.1080/03075070600572132

Chauke, T. A., Mkhize, T. R., Methi, L., & Dlamini, N. (2024). Postgraduate students' perceptions on the benefits associated with artificial intelligence tools on academic success: A case of ChatGPT AI tool. *Journal of Curriculum Studies Research, 6*(1), 46–64. https://doi.org/10.46303/jcsr.2024.4

Cowling, M., Crawford, J., Allen, K. A., & Wehmeyer, M. (2023). Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research supervision. *Australasian Journal of Educational Technology, 39*(6), 17–32. https://doi.org/10.14742/ajet.8598

Cui, J. (2024). The impact of general artificial intelligence on university students' research and technological innovation: A case study of DeepSeek AI and ChatGPT [Preprint]. *SSRN.* https://doi.org/10.2139/ssrn.5179907

Cullen, C. J., Hertel, J. T., & Nickels, M. (2020). The roles of technology in mathematics education. *The Educational Forum, 84*(2), 166–178. https://doi.org/10.1080/00131725.2020.1698683

Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology, 39*(6), 143–159. https://doi.org/10.14742/ajet.8843

Daramola, O. (2021). Lessons from postgraduate supervision in two African universities: An autoethnographic account. *Education Sciences, 11*(7), Article 345. https://doi.org/10.3390/educsci11070345

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior.* Springer. https://doi.org/10.1007/978-1-4899-2271-7

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification." In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments* (pp. 9–15). ACM. https://doi.org/10.1145/2181037.2181040

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Journal of Educational Technology & Society*, 18(3), 75-88.

Dichev, C., & Dicheva, D. (2017). Gamifying education: What is known, what is believed and what remains uncertain: A critical review. *International Journal of Educational Technology in Higher Education*, 14(1), 1–36. https://doi.org/10.1186/s41239-017-0042-5

Dockendorff, M., & Gómez Zaccarelli, F. (2025). Successfully preparing future mathematics teachers for digital technology integration: A literature review. *International Journal of Mathematical Education in Science and Technology*, 56(5), 948–979. https://doi.org/10.1080/0020739x.2024.2309273

Emilsson, U. M., & Johnsson, E. (2007). Supervision of supervisors: On developing supervision in postgraduate education. *Higher Education Research & Development*, 26(2), 163–179. https://doi.org/10.1080/07294360701310797

Ferris, D. (2011). *Treatment of error in second language student writing* (2nd ed.). University of Michigan Press. https://doi.org/10.3998/mpub.2173290

Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. In *Machine learning and the city* (Chapter 45). Wiley. https://doi.org/10.1002/9781119815075.ch45

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies on gamification. *Proceedings of the 47th Hawaii International Conference on System Sciences*, 3025–3034. https://doi.org/10.1109/HICSS.2014.377

Hidayat, A., & Firmanti, P. (2024). Navigating the tech frontier: A systematic review of technology integration in mathematics education. *Cogent Education*, 11(1), Article 2373559. https://doi.org/10.1080/2331186x.2024.2373559

Indriasari, T. D., Luxton-Reilly, A., & Denny, P. (2020). Gamification of student peer review in education: A systematic literature review. *Education and Information Technologies*, 25(6), 5481–5503. https://doi.org/10.1007/s10639-020-10228-x

Jones, B. (2017). Factors in postgraduate supervision that impact on the quality of research at a selected department at a university of technology (Master's dissertation, Durban University of Technology). DUT Institutional Repository. https://doi.org/10.51415/10321/1429

Kızıltaş, Y. (2025). Integration of artificial intelligence (AI) into primary school students' writing skills: The impact of ChatGPT on creative writing and writing self-efficacy. *Journal of Educational Computing Research. Advance online publication.*https://doi.org/10.1177/07356331251365187

Landers, R. N. (2014). Developing a theory of gamified learning: Linking serious games and gamification of learning. *Simulation & Gaming*, 45(6), 752–768. https://doi.org/10.1177/1046878114563660

Laurillard, D. (2013). *Rethinking university teaching: A conversational framework for the effective use of learning technologies* (2nd ed.). Routledge.

Lathwesen, C., & Belova, N. (2021). Escape rooms in STEM teaching and learning—Prospective field or declining trend? A literature review. *Education Sciences*, 11(12), 788. https://doi.org/10.3390/educsci11060308

Leon, J., Núñez, J. L., & Liew, J. (2015). Self-determination and STEM education: Effects of autonomy, motivation, and self-regulated learning on academic achievement. *Educational Psychology*, 35(3), 328–336. https://doi.org/10.1016/j.lindif.2015.08.017

Li, L., Hew, K. F., & Du, J. (2024). Gamification enhances student intrinsic motivation, perceptions of autonomy and relatedness, but minimal impact on competency: A meta-analysis and systematic review. *Educational Technology Research and Development*, 72(2), 1561–1590. https://doi.org/10.1007/s11423-023-10337-7

Looyestyn, J., Kernot, J., Boshoff, K., Ryan, J., Edney, S., & Maher, C. (2017). Does gamification increase engagement with online programs? A systematic review. *PLOS ONE, 12(3), e0173403. https://doi.org/10.1371/journal.pone.0173403*

Luckin, R. (2024). Nurturing human intelligence in the age of AI: Rethinking education for the future. *Development and Learning in Organizations:An International Journal, 1*(4), 1–4. https://doi.org/10.1108/DLO-04-2024-0108

Majuri, J., Koivisto, J., & Hamari, J. (2018). Gamification of education and learning: A review of empirical literature. In *Proceedings of the 2nd International GamiFIN Conference* (pp. 11–19).

Makokotlela, M. V. (2024). Digitising the supervision of postgraduate students in higher education. In D. Guralnick, M. E. Auer, & A. Poce (Eds.), *Advances in higher education and professional development* (Chapter 7). IGI Global. https://doi.org/10.4018/979-8-3693-1289-6.ch007

Maistry, S. (2015). Towards a humanising pedagogy: An autoethnographic reflection of my emerging postgraduate research supervision practice. *Journal of Education*, 61, 35–52. https://doi.org/10.17159/i62a05

Marengo, A., & Pange, P. J. (2024). Envisioning general AI in higher education: Transforming learning paradigms and pedagogies. In D. Guralnick, M. E. Auer, & A. Poce (Eds.), *Creative approaches to technology-enhanced learning for the workplace and higher education* (Vol. 1150, pp. 28–41). Springer. https://doi.org/10.1007/978-3-031-72430-5_28

Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Luansi Ero, P. E., Marlini, C., Zefrin, Z., & Al Fuad, Z. (2025). Critical thinking in the age of AI: A systematic review of AI's effects on higher education. *Education and Information Technologies*, 14(1), Article e2025031. https://doi.org/10.22521/edupij.2025.14.31

Muchuweni, T., Jojo, Z., & Kariyana, I. (2025). Enhancing mathematics instruction through Quizizz: A systematic literature review. *International Journal of Learning, Teaching and Educational Research*, 24(10), 106–124. https://doi.org/10.26803/ijlter.24.10.5

Müller, C., & Mildenberger, T. (2021). Facilitating flexible learning by replacing classroom time with an online learning environment: A systematic review of blended learning in higher education. *Educational Research Review*, 34, 100394. https://doi.org/10.1016/j.edurev.2021.100394

Ng, B. B., Liu, W. C., & Wang, J. C. K. (2012). Self-determination and classroom engagement of adolescents in Singapore. *Learning and Individual Differences*, 22(6), 561–566. https://doi.org/10.1016/j.lindif.2012.07.004

Nicholson, S. (2015). A recipe for meaningful gamification. In *Gamification in education and business* (pp. 1–20). Springer. https://doi.org/10.1007/978-3-319-10208-5_1

Olsher, S., Yerushalmy, M., & Chazan, D. (2016). How might the use of technology in formative assessment support changes in mathematics teaching? *For the Learning of Mathematics, 36*(3), 11–18. https://www.jstor.org/stable/44382716

Omodan, B. I., & Marongwe, N. (2024). The role of artificial intelligence in decolonising academic writing for inclusive knowledge production. *Interdisciplinary Journal of Education Research, 6*(S1), 1–14. https://doi.org/10.38140/ijer-2024.vol6.s1.06

Oparinde, K. (2021). Postgraduate supervision: A heuristic approach to learning, unlearning, and relearning. *Asian Journal of University Education, 17*(4), 243–250. https://doi.org/10.24191/ajue.v17i4.16202

Ortiz-Rojas, M., Chiluiza, K., Valcke, M., & Bolanos-Mendoza, C. (2025). How gamification boosts learning in STEM higher education: A mixed methods study. *International Journal of STEM Education, 12*(1), Article 73. https://doi.org/10.1186/s40594-024-00521-3

Oulaich, S. (2019). Pedagogy in the digital age. In *Proceedings of the 4th International Conference on E-Learning and Games (ICELG '19)* (pp. 1–6). ACM Digital Library. https://doi.org/10.1145/3368756.3369058

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, **372**, n71. https://doi.org/10.1136/bmj.n71

Qudsi, H. (2024). Gamification in education: Boosting student engagement and learning outcomes. *ShodhKosh: Journal of Visual and Performing Arts, 5*(4), 686–693. https://doi.org/10.29121/shodhkosh.v5.i4.2024.2542

Rahmi, I., Rimenda, T., & Ariyanti, T. D. (2025). Gamification as an alternative to increase students' motivation: A scoping review. *Journal of Education and Learning (EduLearn), 19*(2), 233–241. https://doi.org/10.11591/edulearn.v19i2.21771

Ramadhan, A., Warnars, H. L. H. S., & Abdul Razak, F. H. (2024). Combining intelligent tutoring systems and gamification: A systematic literature review. *Education and Information Technologies, 29*(1), 77–98. https://doi.org/10.1007/s10639-023-12092-x

Reeve, J. (2012). A self-determination theory perspective on student engagement. In S. Christenson, A. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 149–172). Springer. https://doi.org/10.1007/978-1-4614-2018-7_7

Reeve, J. (2023). Cognitive evaluation theory. In *The Oxford handbook of self-determination theory* (Chapter 3). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197600047.013.3

Rivera, E. S. (2021). Gamification for student engagement: A framework. *Journal of Further and Higher Education*, 45(7), 999–1012. https://doi.org/10.1080/0309877X.2021.1875201

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behaviour. *Psychological Inquiry*, 11(4), 227–268. https://doi.org/10.1207/s15327965pli1104_01

Ryan, R. M., & Deci, E. L. (Eds.). (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press. https://doi.org/10.1521/978.14625/28806

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review*, 32(1), 77–112. https://doi.org/10.1007/s10648-019-09498-w

Sánchez-Arévalo, B. M., Valenciano-Valcárcel, J., & Fernández-Cézar, R. (2025). Analysis on the use of gamification in mathematics in primary education: A literature review. *Mathematics Education Research Journal*, 37(1), 30–55. https://doi.org/10.17583/redimat.16133

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies*, 74, 14–31. https://doi.org/10.1016/j.ijhcs.2014.09.006

Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. John Wiley & Sons.

Shumba, T. (2024). Exploring lecturers' readiness and perceptions of gamification in higher education institutions of South Africa. *Journal of Public Administration and Development Alternatives*, 9(3), 115–127. https://doi.org/10.55190/jpada.2024.346

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. https://doi.org/10.3102/0034654307313795

Sim, K. N., Northcote, M., & Lim, C. P. (2023). Technology-enabled undergraduate and postgraduate research supervision. *Australasian Journal of Educational Technology*, 39(6), 33–48. https://doi.org/10.14742/ajet.9149

Standage, M. (2023). Self-determination theory applied to sport. In *The Oxford handbook of self-determination theory* (Chapter 35). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197600047.013.35

Subhash, S., & Cudney, E. A. (2018). Gamified learning in higher education: A systematic review of the literature. *Computers in Human Behavior*, 87, 192–206. https://doi.org/10.1016/j.chb.2018.05.028

Triantafyllou, S. A., Georgiadis, C., & Sapounidis, T. (2025). Gamification in education and training: A literature review. *International Review of Education*. https://doi.org/10.1007/s11159-025-10034-z

Veletsianos, G., & Kimmons, R. (2012). Networked participatory scholarship: Emergent techno-cultural pressures toward open and digital scholarship in online networks. *Computers & Education*, 58(2), 766–774. https://doi.org/10.1016/j.compedu.2011.10.001

Venter, M., & de Wet, L. (2024). The influence of player type on the motivation of students in a gamified programming learning environment. *ICT in Education Journal*, 19(2), 88–102. https://doi.org/10.1007/978-3-031-82478-4_1

Viberg, O., Grönlund, Å., & Andersson, A. (2020). Integrating digital technology in mathematics education: A Swedish case study. *Interactive Learning Environments*, 31(1), 232–243. https://doi.org/10.1080/10494820.2020.1770801

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.

Wingate, U. (2006). Doing away with "study skills." *Teaching in Higher Education*, 11(4), 457–469. https://doi.org/10.1080/13562510600874268

Willcoxson, L. (1998). The impact of academics' learning and teaching preferences on their teaching practices: A pilot study. *Studies in Higher Education*, 23(1), 59–70. https://doi.org/10.1080/03075079812331380492

Winkler, R., & Söllner, M. (2018). Unleashing the potential of chatbots in education: A state-of-the-art analysis. *Academy of Management Proceedings*, 2018(1), Article 15903. https://doi.org/10.5465/ambpp.2018.15903

Xu, C. (2025). Research on security risks and supervision of generative AI large models. In *Proceedings of the 2025 International Conference on Artificial Intelligence and Educational Systems* (pp. 120–128). ACM. https://doi.org/10.1145/3744367.3744395

Yığ, K. G., & Sezgin, S. (2021). An exploratory holistic analysis of digital gamification in mathematics education. *Journal of Educational Technology and Online Learning, 4*(2), 115–136. https://doi.org/10.31681/jetol.888096

Yllana-Prieto, F., González-Gómez, D., & Jeong, J. S. (2025). The escape room and breakout as an aid to learning STEM contents in primary schools: An examination of the development of pre-service teachers in Spain. *Education* 3-13, 53(1), 43–57. https://doi.org/10.1080/03004279.2022.2163183

Zainuddin, Z., Chu, S. K. W., & Perera, C. J. (2024). Gamification platforms for flipped learning implementation. In Z. Zainuddin, S. K. W. Chu, & C. J. Perera (Eds.), *Gamification in a flipped classroom* (pp. 45–62). Springer. https://doi.org/10.1007/978-981-97-2219-8_5

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Recommendations for future research. *International Journal of Educational Technology in Higher Education, 16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

CHAPTER TEN

# Sustainable Mentorship Practices: Designing Frameworks That Blend AI Efficiencies with Human Intuition

**Adelaide R. M. Pitso**[1]

AFFILIATIONS

[1]Department of Tourism, Hospitality and Sport Management, Faculty of Management and Public Administration Sciences, Walter Sisulu University, Butterworth, South Africa.

CORRESPONDENCE

Email: ambili@wsu.ac.za

REFERENCE

Pitso, A. R. M. (2026). Sustainable mentorship practices: Designing frameworks that blend AI efficiencies with human intuition. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 166–190). ERRCD Forum. https://doi.org/10.38140/obp4-2026-10



**Abstract:** The integration of Artificial Intelligence (AI) into academic mentoring has the potential to transform conventional approaches by increasing efficiency, availability, and accuracy. However, the challenge lies in ensuring that such advances do not supplant the essential human factors of empathy, intuition, and contextual understanding. This chapter aims to design a sustainable framework that combines AI-driven effectiveness with human-centred mentorship practices, thereby achieving an optimal equilibrium between technological advancement and personalised guidance. To this end, a qualitative methodological approach was employed, with data collected through semi-structured interviews with ten academic mentors and fifteen postgraduate mentees from a range of multidisciplinary fields of study. Thematic data analysis was utilised to examine the data. The findings reveal that while AI significantly enhances routine tasks such as feedback, scheduling, and resource allocation, mentees consistently value human interaction for emotional support, nuanced advice, and contextual adaptability. The proposed framework underscores the integration of AI tools and human guidance, highlighting areas where AI excels and domains where human insight remains irreplaceable. This chapter emphasises the importance of promoting synergy between AI and human mentors, which ultimately aims to improve the quality, accessibility, and inclusiveness of mentorship in academia. Furthermore, the chapter serves as a foundation for further research on sustainable, AI-enhanced mentorship paradigms.

***Keywords:*** Artificial intelligence, higher education, mentorship in higher education, postgraduate supervision, university sustainability.

## 1. Introduction

The landscape of academic mentoring is changing, and higher education institutions aim to boost productivity while embracing the diverse backgrounds of postgraduate mentees and providing support (Khumalo & Ndlova, 2024). Akinwalere and Ivanov (2022) noted that as institutions strive to enhance productivity and create fairer, more inclusive educational environments, new technologies, such as Artificial Intelligence (AI), are making significant impacts. However, these technologies also introduce serious ethical questions and complexities. Traditionally, mentorship has been grounded in knowledge transfer, engagement, and empathy; yet, it now finds itself caught between tradition and innovation (Kuznetsova-Bogdanovitsh & Jyrämä, 2022). AI technologies promise to enhance efficiency in various ways, from automating everyday communications and feedback to providing personalised learning suggestions and predictive analytics (Holmes et al., 2021). Nevertheless, there is a risk that AI may overlook the critical emotional and intuitive components essential for fostering academic and professional development (Mohamed Badawy et al., 2023). Therefore, it is crucial to explore how a sustainable mentorship framework can seamlessly integrate AI capabilities with human intuition to improve the mentoring experience while maintaining its core integrity.

According to Sarany et al. (2023), mentoring is a cornerstone of success in the ever-evolving academic environment, particularly at the postgraduate level, where mentees navigate complex research requirements and personal development goals. Zellers et al. (2008) highlighted that traditional mentorship models often involve face-to-face interactions, informal knowledge exchange, and adaptive coaching based on shared academic and cultural backgrounds. Despite their demonstrated effectiveness, these models are frequently constrained by mentor availability, scalability, and varying levels of participation and quality. Consequently, AI-powered systems have emerged to help bridge these gaps by analysing mentee performance, providing scalable solutions for rapid feedback, and predicting academic risks (Luckin et al., 2016). Nevertheless, their use raises significant concerns regarding the validity of the mentor-mentee relationship, particularly regarding emotional intelligence, cultural sensitivity, and ethical responsibility. This creates a major question: "How can higher education institutions utilise AI while preserving the relational depth of effective mentorship?" Existing mentorship models often operate in isolation, either being overly reliant on human connection or prematurely adopting technology without appropriate ethical and pedagogical considerations. As a result, mentees may encounter diminished personal relationships, fragmented support, and unequal outcomes. Additionally, an overreliance on AI risks depersonalisation, cultural insensitivity, and algorithmic bias, particularly in diverse academic settings.

Kaye (2024) stated that human intuition and empathy are important in mentoring situations involving underrepresented first-generation or international mentees. These mentees often rely extensively on mentors for socio-economic support, contextual navigation, and specialised academic help. Although AI can mimic specific patterns through natural language processing (NLP), it lacks the potential for profound human connection and moral judgement (Marcus & Davis, 2019). Furthermore, reliance on AI tools might reduce mentorship to transactional exchanges devoid of relational richness. Consequently, a hybrid strategy that preserves the qualities of human mentorship while harnessing AI capabilities is essential for establishing sustainable and equitable mentorship practices in academia.

In higher education, the urgency for sustainable mentorship models has been intensified by global trends such as resource constraints, academic burnout, and remote learning (Essop, 2021). Hence, Bao (2020) noted that the COVID-19 pandemic exposed the vulnerability of traditional mentoring systems and catalysed interest in digital tools for resilience and continuity. In this context, AI-enabled mentoring platforms may provide adaptive learning environments, resources, and real-time progress tracking across institutional and geographical boundaries

(Imamguluyev et al., 2024). Nevertheless, mentorship sustainability must encompass operational efficiency, continuity, adaptability, and relational and long-term impact (Stozhko et al., 2021). Thus, developing a framework that systematically integrates AI's precision with the human mentor's intuition and ethical judgment is both timely and important.

## 2. Literature Review

### 2.1 The changing landscape of academic mentorship in higher education

Academic mentorship has long been an important part of student development, particularly at the postgraduate level (Stravakou & Lozgka, 2022). Traditionally, mentorship relied on in-person interactions and informal relationships in which senior academics assisted junior scholars with research, professional development, and career paths (Nuis et al., 2023). According to Jyoti and Sharma (2015), these relationships were often based on physical proximity, shared institutional culture, and a mutual interest in academic achievement. While traditional mentorship models fostered strong interpersonal bonds, they were heavily reliant on the availability and willingness of individual mentors, resulting in inconsistencies in quality and accessibility (Nuis et al., 2023). Furthermore, the historical foundations of academic mentorship were established in the apprenticeship model, in which the mentee gradually acquired tacit knowledge and professional skills through close, long-term interactions with a mentor (Peiser et al., 2018). This model prioritised academic guidance along with the development of personal and ethical values. However, as higher education systems have expanded and diversified, the demand for mentorship has increased, often outpacing institutional capacity. The modern academic landscape currently faces various significant challenges, particularly in postgraduate supervision, such as limited faculty time, mentor burnout, and high student-to-supervisor ratios, which have led to fragmented mentoring experiences (Sambunjak et al., 2010; Odularu & Akande, 2024).

The acceleration of digital transformation in education, particularly after the COVID-19 pandemic, has fuelled a trend towards remote and technology-enhanced mentoring (Bao, 2020). For example, virtual mentorship provides flexibility and increased access, particularly for students in remote or underserved areas. However, it also raises concerns regarding relational depth and engagement quality (Owen, 2015). Iglesias-Pradas et al. (2021) noted that remote models decrease logistical barriers to regular engagement, but they often fail to replicate the complex, spontaneous exchanges that occur in person. Furthermore, remote mentorship may exacerbate inequities in areas where students lack a reliable internet connection, digital literacy, or safe study environments.

In addition to technological advancements, there is an increasing emphasis on the sustainability of mentorship practices. Thus, Stozhko et al. (2021) highlighted that mentorship must now be designed to meet the requirements of a diverse and changing student population without overburdening academic staff. Furthermore, sustainable mentorship systems provide continuity, inclusivity, and scalability. As a result, higher education institutions must reconsider how mentorship is structured, evaluated, and supported across disciplines (Lechuga, 2011). According to Coetzee (2023), the old "one-size-fits-all" strategy is increasingly seen as inadequate for meeting the diverse demands of students from various socio-economic and cultural backgrounds, particularly in globalised academic institutions.

Equity concerns are prevalent in mentorship literature. Hagler (2023) indicated that marginalised and first-generation students frequently receive less consistent and culturally appropriate mentorship, and these inequities may be exacerbated by unconscious biases or institutional attitudes that exclude non-traditional student experiences. As mentorship models expand, it is critical that frameworks include the lived realities of different students and consciously incorporate equality into mentor training and institutional design (Ntshongwana, 2024).

Literature reveals a significant shift in academic mentorship, driven by structural constraints and digital innovation. The current changes require a purposeful and hybridised strategy that preserves the relational nature of mentorship while embracing the scale and flexibility technology provides. Furthermore, this shift opens the door for frameworks that integrate AI-enabled efficiencies with human-centred principles to promote a more sustainable and inclusive future for academic mentorship.

### 2.2 Artificial intelligence in mentorship: Tools, capabilities, and constraints

Integrating AI into higher education has transformed traditional academic practices, including mentorship (Akinwalere & Ivanov, 2022). AI is increasingly being used to improve the efficiency, scalability, and accessibility of mentoring systems, particularly at the postgraduate level. The authors state that AI technologies such as chatbots, predictive analytics, automated feedback tools, and scheduling algorithms have been developed to meet the growing demand for academic assistance (Holmes et al., 2021; Akinwalere & Ivanov, 2022). Automated feedback systems are among the most commonly used AI tools in education (Fitria, 2021). Saha and Mondal (2024) mention that these systems analyse student submissions and deliver immediate, personalised responses that help students identify errors and enhance their performance. Furthermore, AI-powered chatbots and virtual assistants are utilised to answer routine questions, assist students with academic procedures, and provide information about institutional services (Zawacki-Richter et al., 2019). AI technologies reduce administrative duties for academic staff, allowing them more time for one-on-one human interactions. Moreover, AI-enhanced scheduling solutions, such as intelligent calendar integrations, optimise meeting times between mentors and mentees based on availability, eliminating inefficiencies associated with back-and-forth coordination (Suman et al., 2024). Additionally, performance analytics enabled by machine learning provide valuable insights into student progress, engagement levels, and potential risk indicators. These insights allow mentors to proactively identify mentees who require intervention and tailor support techniques accordingly (Ncube & Ngulube, 2024; Alalawi et al., 2024). This means that AI helps provide more targeted and data-driven mentorship, aligning with the increasing emphasis on evidence-based techniques in higher education. Despite the benefits of AI, there are limits to its efficacy in mentoring. One of the most pressing concerns is that AI systems lack emotional intelligence, intuition, and cultural sensitivity (Selwyn, 2019).

Effective mentoring involves an understanding of complex social settings, nuanced communication skills, and human emotions, all of which AI cannot yet replicate. This implies that AI cannot replace humans' ability to empathise and make ethical decisions. Additionally, overreliance on AI poses both ethical and technical challenges. Algorithmic bias is a well-documented issue where AI systems trained on non-representative data may unintentionally promote social inequities (Binns, 2018). For example, predictive algorithms used in student assessment and feedback systems may favour dominant cultural perspectives while disadvantaging students from marginalised backgrounds. Tsai et al. (2020) noted that in disciplines reliant on collaborative reflection and dialogue, the depersonalisation of mentorship exchanges through AI interfaces might weaken mentees' sense of belonging and trust. Williamson and Eynon (2020) highlighted another important concern: the transparency and accountability of AI systems. In mentorship, mentees may not be aware of how their data is collected, processed, or used to support AI-driven decisions. Moreover, a lack of transparency may lead to mistrust, especially if a mentee feels surveilled or judged by impersonal algorithms. This underscores that the ethical use of AI tools requires careful consideration of consent, data governance, and human monitoring. According to Ajani et al. (2024) and Imamguluyev et al. (2024), AI has developed promising tools to support and streamline various aspects of the mentorship process in higher education. While these tools improve operational efficiency and provide actionable information, they must be employed judiciously to ensure they complement human-centred approaches that uphold the relational and ethical foundations of mentorship. Therefore, AI should not replace human mentors but rather supplement their efforts with intelligent tools that enhance access and quality without compromising personalisation or trust.

### 2.3 Human-centred mentorship: Empathy, intuition and relational depth

Despite the growing integration of AI into academic processes, human-centred mentorship remains a critical and irreplaceable component of successful postgraduate education (Chan & Tsi, 2023). Mentorship is primarily a relationship that fosters professional development, emotional resilience, and intellectual identity. Lechuga (2011) noted that the most impactful mentors provide more than just technical assistance; they also offer empathy, encouragement, and contextualised support—qualities that existing AI systems cannot meaningfully replicate. Deane et al. (2022) indicated that empathy and emotional support are highly significant in mentorship because they help mentees feel understood and valued, especially during difficult periods of academic work. In agreement, NASEM (2019) stated that mentors who offer reassurance, validation, and emotional regulation are more likely to foster mentee confidence and persistence. Johnson and Ridley (2018) observed that trust built through consistency and sincere human interaction enables mentees to share vulnerabilities, engage in reflective practice, and ask critical questions without fear of judgement. This trust forms the foundation for what is commonly referred to as "psychosocial support," a component of mentorship that is critical to mentee success and well-being.

Cultural responsiveness and context sensitivity are essential in a diverse and multicultural academic environment. This implies that good academic mentors understand and respect their mentees' backgrounds, values, and life experiences, modifying their support strategies accordingly (Straus et al., 2013). Sachpasidi et al. (2024) concurred that culturally sensitive mentoring helps to reduce marginalisation, especially among mentees from historically disadvantaged backgrounds. Unlike AI, which frequently uses biased data sets, human mentors can navigate complex social dynamics and tailor their approaches to match their mentees' cultural and institutional contexts (Farrelly & Baker, 2023). From the mentee's perspective, personalised and relational mentorship is not only preferable but often necessary. Curtin et al. (2016) and Stravakou and Lozgka (2022) found that mentees prefer mentors who understand their academic journeys, career goals, and personal challenges. These insights stem not only from verbal communication but also from mentors' ability to recognise subtle clues and respond intuitively, a skill AI cannot replicate. In a study conducted by Singe et al. (2021), doctoral students emphasised that the most important mentorship experiences were those in which the mentor interacted with them holistically, taking into account both academic and personal dimensions of growth.

Similarly, Lechuga (2011) found that mentors who shared their academic challenges helped normalise failure and reduce performance anxiety in mentees, which is another quality that would be difficult to reproduce with AI. These studies demonstrate that AI cannot replace the nature of human mentorship; mentees reported higher satisfaction and retention when mentors displayed empathy and discussed their own experiences (Nuis et al., 2023). While AI may be useful for administrative or analytical tasks, it lacks the emotional intelligence and relational depth needed for transformative mentorship (Dwivedi, 2025). Furthermore, human intuition allows mentors to recognise unspoken difficulties, respond sympathetically to discomfort, and empower students in complex socio-emotional situations. Rather than seeking to replace human mentorship, institutions should enhance these human-centric features while allowing AI to serve as a helpful operational tool.

### 2.4 Toward a hybrid framework: Merging AI efficiencies with human intuition for sustainable mentorship

The continuing digital transformation of higher education has paved the way for innovative mentorship models that combine AI and human-centred practices. According to Sajja et al. (2024), this hybrid approach takes advantage of AI systems' efficiency while maintaining the empathy, intuition, and contextual awareness that only human mentors can provide. Theoretical foundations, such as socio-technical systems theory, which proposes the necessity for alignment between social actors and technological instruments, provide a guiding framework for this integration (Gumede & Tladi, 2023). Additionally, socio-technical perspectives encourage a balance between human agency and machine functionality, acknowledging that neither operates effectively in isolation. In the context of academic mentorship, this means using AI to manage repetitive administrative tasks, such as scheduling or feedback, while reserving emotionally nuanced interactions and decision-making for human mentors (Ciriello et al., 2024). The literature suggests that such a blended approach can increase the scalability and responsiveness of mentorship without sacrificing relational depth (Holmes et al., 2021).

Several frameworks and pilot programmes are currently testing this integration. For example, Georgia Tech uses Jill Watson, an AI teaching assistant built on IBM's Watson platform. This tool demonstrates how AI can help human educators by managing frequently asked student questions and freeing time for greater mentor-mentee connection (Goel & Polepeddi, 2016). Similarly, hybrid mentoring systems, such as those launched at Mohammed VI Polytechnic University, combine AI-driven predictive analytics with human oversight, allowing mentors to provide personalised guidance based on real-time student data (Baba et al., 2024; Goel et al., 2016). However, Williamson and Eynon (2020) noted that the effective implementation of hybrid mentorship models depends on inclusive design and ethical principles; as AI tools gain traction in education, concerns about data privacy, algorithmic bias, and transparency become critical. Furthermore, inclusive mentorship systems must be created with various mentee populations in mind, ensuring that AI does not exacerbate current inequalities but promotes fairer access to assistance and resources (Binns, 2018). Sustainability and scalability are critical parameters for assessing the efficacy of blended mentorship programmes. Zawacki-Richter et al. (2019) found that AI-supported mentorship can help higher education institutions with limited workforces to satisfy increasing student demand, particularly in postgraduate supervision. Slimi (2023) highlighted that AI improves mentoring delivery efficiency by automating mundane processes and making certain functions available around the clock. Nonetheless, experts caution against over-automation and stress that mentorship's long-term influence depends on developing relationships, trust, and social learning domains where human engagement is still required (Selwyn, 2019).

## 3. Problem Statement

Literature has shown that, with increased effectiveness, scalability, and data-driven support, the rapid adoption of AI in higher education has opened up new avenues for improving academic mentoring (Luckin et al., 2016; Donner & Hummel, 2025). However, these technological developments have also raised concerns about the potential decline of human traits such as empathy, intuition, contextual awareness, and cultural sensitivity, all of which are essential for successful mentoring (Rachmad, 2022; Shen et al., 2024; Singh & Singh, 2025). Despite being rich in relationships, Nkoala (2024) and Mbanjwa (2025) noted that traditional mentorship models are becoming increasingly strained due to staffing shortages, a rise in postgraduate enrolment, and increased student diversity. Conversely, new AI-driven technologies often lack the contextual awareness, emotional nuance, and moral discernment necessary to foster comprehensive postgraduate growth (Marcus & Davis, 2019). Higher education institutions are now faced with an increasing dilemma: how to effectively incorporate AI tools into mentoring without sacrificing the ethical integrity and depth of human guidance. There is currently no comprehensive or long-lasting framework that intentionally combines AI efficiencies with human-centred mentorship methods, despite ongoing testing with AI-supported mentoring systems. In the absence of such a framework, mentoring encounters risk being disjointed, unfair, or overly automated, particularly for disadvantaged or underrepresented postgraduate groups (Judijanto, 2025). Therefore, the necessity of creating a sustainable, morally sound hybrid mentorship model that capitalises on AI's advantages while maintaining the indispensable human components required for significant academic support is the main issue this study seeks to address. Consequently, the following research questions need to be addressed in order to develop this framework:

- How can Artificial Intelligence tools be integrated into academic mentorship to enhance efficiency while preserving the essential human factors of empathy, intuition, and contextual relevance?
- What are the key challenges and benefits experienced by academic mentors and mentees when engaging with hybrid AI-human mentorship models?
- How can a sustainable framework for AI-supported mentorship be developed to ensure inclusivity, scalability, and ethical practices in diverse academic environments?

This study contributes to the emerging discourse on ethical AI integration in education by adopting a pragmatic and inclusive approach to mentorship. Furthermore, it offers actionable insights for lecturers, policymakers, and educational technologists interested in designing scalable, empathetic, and future-proof academic support systems. As higher education institutions grapple with the best ways to integrate AI into pedagogical settings, this chapter serves as both a theoretical foundation and a practical guide for designing a mentorship framework that is not only technologically advanced but also emotionally intelligent and culturally responsive.

## 4. Materials and Methods

This study employed a qualitative research approach, which was suitable for generating in-depth and contextually grounded insights into how AI can be integrated with human-centred mentorship practices in higher education. The aim of the study was not to test hypotheses or measure variables but rather to conceptualise and develop a sustainable hybrid mentorship framework. A qualitative orientation enabled the exploration of the lived experiences, perceptions, and expectations of mentors and mentees. This approach supported the identification of patterns, tensions, and opportunities within current mentorship practices and

AI-assisted tools, reflecting the socially constructed and interpretive nature of mentorship in higher education. It aligns with contemporary qualitative methodologies used in higher education studies, where evolving technologies intersect with complex interpersonal academic processes.

A descriptive study design was employed to inform the methodical development of a sustainable AI-enhanced mentorship framework. This approach was deemed acceptable as it allowed the researcher to collect and interpret participants' opinions in a manner that remained true to their actual words, while also facilitating the conceptual integration of insights from interviews and literature. Since the study does not aim to provide empirical generalisations, the design prioritises depth, contextual understanding, and interpretive synthesis.

The study targeted two broad populations within one university in the Eastern Cape such as all academics across faculties who are involved in postgraduate supervision and have exposure to traditional or AI-supported mentorship practices. Additionally, all registered Master's and PhD students across disciplines who have engaged in either in-person or AI-assisted academic mentorship were also targeted. From these populations, a purposive sample was selected to ensure participants had direct experience relevant to academic mentorship. The final sample consisted of 10 academic mentors drawn from 6 different faculties (Economics and Finance Science, Education, Engineering, Law, Humanities and Social Science, Management and Public Administration Sciences, and Medicine and Health Sciences) and were selected based on their postgraduate supervision experience and familiarity with mentorship tools (traditional and AI-supported). Furthermore, 15 postgraduate mentees (Master's and PhD) across the 6 faculties were chosen for their experience with either in-person or AI-assisted mentorship. Participants were selected to ensure diversity in discipline, gender, and institutional background to enhance the depth and transferability of findings.

The study employed semi-structured interviews because they allow participants to share rich and in-depth insights while still ensuring consistency across key topics. This flexibility made it possible to explore diverse experiences and perceptions of AI-supported mentorship in a structured yet adaptive manner. The interview protocol was developed based on themes from the literature review, focusing on the following: Firstly, Theme 1: Experiences with traditional mentorship. This theme corresponds to understanding the current context of academic mentorship, which informs the research question on how AI tools can be integrated without compromising human-centred elements such as empathy, intuition, and contextual relevance. Secondly, Theme 2: Exposure to or opinions on AI-assisted mentorship tools. This theme directly addresses participants' experiences with AI technologies, supporting the identification of key benefits and challenges associated with hybrid AI-human mentorship models. Furthermore, Theme 3: Perceptions of empathy, trust, and emotional support in academic supervision. This theme specifically investigates the human factors critical to mentorship, enabling exploration of how AI integration might preserve or affect these relational elements.

Lastly, Theme 4: Ideal features of a hybrid AI-human mentorship model. This theme aligns with the research question on developing a sustainable framework by gathering participants' perspectives on inclusive, scalable, and ethically grounded mentorship practices that integrate AI. Each interview lasted 30–45 minutes and was conducted virtually on MS Teams or face-to-face, depending on the participant's preference and location. Interviews were audio-recorded with consent and later transcribed verbatim for analysis.

Participants were recruited through university academic networks, supervisor lists, and postgraduate student mailing lists. An invitation letter was distributed, which included details about the study, such as its purpose, significance, and ethical considerations. As a result, interested individuals contacted the researchers directly and provided informed consent before participating. Participants were informed that their involvement was voluntary and that they could withdraw at any stage.

### 4.1 Data analysis

The data were analysed using NVivo version 15, following Braun and Clarke's (2006) six-phase thematic analysis process. This process involved familiarising oneself with the data, generating initial codes, searching for potential themes, reviewing and refining those themes, defining and naming them, and finally producing the analytical report. The resulting themes were closely aligned with the study's objectives and existing literature, allowing for the identification of areas of convergence and divergence between AI capabilities and human mentorship practices. During the semi-structured interviews, saturation became evident when no new themes, perspectives, or insights emerged from additional participants, and the data began to repeat across both mentor and mentee groups. By the time the final interviews were conducted, the patterns were consistent, and further data collection was unlikely to add new conceptual information. Saturation, therefore, confirmed that the sample size was sufficient for addressing the study's aims and developing the proposed hybrid mentorship framework.

### 4.2 Ethical considerations

Ethical clearance was obtained from the Research Ethics Committee of Walter Sisulu University, protocol number 31/16/10/2025/DRI. Participants were informed of their rights, which included the right to anonymity, confidentiality, and voluntary participation. Data was securely stored on password-protected systems, and pseudonyms were used in reporting to protect identities. No personal information or identifiers were linked to the published data. Furthermore, this study adhered to South African academic ethical guidelines and the Protection of Personal Information Act (POPIA).

## 5. Presentation of Results

This section presents the study's findings, which aim to explore and design a sustainable mentorship framework that effectively integrates the efficiencies of AI with the relational depth of human intuition. The findings from semi-structured interviews with academic mentors and postgraduate mentees, are as follows:

**5.1 Theme 1: Integrating AI tools into academic mentorship**

Across all faculties, both academic mentors and postgraduate students expressed strong appreciation for traditional face-to-face mentorship. Participants emphasised that the human connection remains central to meaningful academic support. When asked the guiding question, "How can AI tools be integrated into academic mentorship to enhance efficiency while preserving essential human factors?", responses consistently highlighted the depth and authenticity of human interaction.

A mentor from the Faculty of Law, Humanities and Social Science shared:

> *"With the mentorship style I use, it allows me to truly know the person behind the student or work. It's more than academic support; it's guiding them through personal, emotional, and career-related issues. That relationship is impossible to replace. Also, I can be able to see when a student is tired, stressed, or losing confidence. AI will never pick up on those subtle cues."*

Another mentor from the Faculty of Engineering added:

> *"Traditional mentorship allows you to intervene early. In face-to-face sessions, I can immediately sense when a student is confused even if they say they are fine."*

However, many mentors described the intensity and strain associated with traditional approaches. A supervisor from the Faculty of Medicine and Health Sciences explained:

> *"The workload is massive. I supervise undergraduates, Masters, and PhDs. I often wish I had support systems that could help with routine questions so I could focus on conceptual guidance. Mentorship becomes overwhelming because supervising multiple students while managing teaching and research is difficult. Sometimes students expect immediate feedback, and I simply cannot respond quickly to everyone."*

A Masters students expressed similar challenges. A postgraduate student from the Faculty of Management and Public Administration Sciences said:

> *"Face-to-face mentorship is powerful, but inconsistent. Sometimes weeks pass before I get proper feedback."*

Another PhD student from the Faculty of Education added:

> *"My supervisor is excellent, but getting time with them is difficult. When they finally respond, it's very helpful but waiting can be stressful.*

A Masters student from the Faculty of Law, Humanities and Social *Science* highlighted emotional dependence on traditional mentorship:

> *"When I am anxious about my progress, I prefer speaking to my mentor directly because they reassure me. AI cannot do that."*

A student from the Faculty of Medicine and Health Science in a Masters programme mentioned:

> *"Getting one-on-one sessions is difficult because my supervisor has so many students. Sometimes I feel like my questions are not seen as urgent."*

Across faculties, participants generally converged on the importance of traditional, human-centred mentorship. Mentors from the Law, Humanities and Social Sciences, Engineering, and Medicine and Health Sciences faculties highlighted empathy, personal connection, and emotional awareness as essential components of supervision. Their reflections frequently emphasised the relational depth that arises from face-to-face engagement. However, divergences emerged based on discipline and supervision experience. For example, mentors from the Faculty of Engineering tended to emphasise efficiency and structured problem-solving, stating that traditional mentorship is effective but time-consuming. At the student level, master's students were more vocal about access challenges, noting that limited supervision time slows progress. PhD students, on the other hand, emphasised the need for sustained intellectual engagement and emotional support, highlighting that doctoral work demands a more personalised supervisory approach. This range of views illustrates that while traditional mentorship is valued across the board, the reasons for valuing it differ by department, academic level, and years of supervisory experience.

### 5.2 Theme 2: Challenges and benefits of hybrid AI–human mentorship models

When asked "What are the key challenges and benefits of engaging with hybrid AI–human mentorship models?", participants provided mixed but insightful responses. While mentors welcomed the efficiency that AI brings, they were cautious about its limitations in relational aspects.

A mentor from the Faculty of Economics and Finance Science noted the efficiency benefits:

> *"AI scheduling and quick feedback systems have reduced my administrative workload. I can now use my time for deeper intellectual engagement."*

A mentor from the Faculty of Medicine and Health Sciences described a similar experience:

> *"I rely on AI to flag formatting issues or incomplete references. It saves hours. But I would never trust AI to give conceptual feedback on a research proposal."*

Others were more cautious. A mentor from the Faculty of Law, Humanities and Social Science remarked:

> *"AI gives generic advice. It doesn't understand creative methodologies or practice-based research."*

A mentor from the Faculty of Education noted:

> *"AI tools reduce admin tasks significantly. Automated scheduling and document feedback systems free my time for deeper conversations with students. But AI cannot build rapport or understand student emotions."*

Students also expressed varied opinions. A master's student from the Faculty of Engineering said:

> *"AI tools help me fix grammar and structure, but when I ask for conceptual clarity, it sometimes gives misleading information."*

A PhD student from the Faculty of Law, Humanities and Social *Science* added:

> *"AI helped me brainstorm ideas, but the suggestions sometimes contradict my supervisor's expectations. That causes confusion. I also use AI systems for quick feedback on structure and grammar, but the responses feel mechanical. It doesn't understand what my argument really means. "*

Some postgraduate students highlighted the advantages of AI for productivity:
*"AI summarised articles for me, which saved a lot of time, especially when I was reviewing many journal papers."* PhD student, Faculty of Law, Humanities and Social Sciences stated.
*"Chatbots helped me prepare for meetings by explaining concepts before I met my supervisor."* PhD student, Faculty of Education noted.
Yet, several students emphasised the risk of over-reliance on AI. Master's student, Faculty of Management Sciences mentioned:

> *"AI made me feel like I was learning less. It gave too many answers, and I stopped thinking critically at some point."*

Some students expressed concerns about over-reliance on AI. A Master's student from the Faculty of Medicine and Health Science warned:

> *"AI made me too dependent at some point. I thought I was improving academically, but I later realised I was just accepting suggestions without understanding them fully."*

Participants across faculties demonstrated a consensus on the practical value of AI tools, particularly for administrative or routine academic tasks. Mentors in the Faculty of Economics and Finance were especially positive about the efficiency of AI in scheduling, formatting checks, automated reminders, and preliminary feedback tools. Mentees in Law, Humanities and Social Sciences, and Medicine and Health Sciences also appreciated AI's ability to provide quick responses and assist with research organisation. However, significant divergences emerged in perceptions of AI's academic reliability. Participants in Law, Humanities and Social Sciences expressed scepticism, arguing that AI-generated feedback lacked nuance and contextual understanding. In contrast, participants in Engineering and Medicine and Health Sciences were more open to using AI for initial technical explanations, grammar checks, and literature structuring. At the student level, PhD candidates were more critical of AI feedback on conceptual and methodological aspects, while Master's students viewed AI as beneficial for building foundational understanding at the early stages. This cross-sectional diversity enhances the understanding of hybrid mentorship models, illustrating that acceptance of AI is influenced by discipline, supervisory experience, and research level.

## 5.3 Theme 3: Developing a sustainable, inclusive, and ethical framework for AI-supported mentorship

### *5.3.1 Sub-theme 1: Perceptions of empathy, trust and emotional support*

Across faculties, participants strongly emphasised that empathy, trust, and emotional support are core pillars of effective mentorship. When asked "How can a sustainable AI-supported mentorship framework ensure inclusivity, scalability, and ethical practice?.", many expressed concerns about the limitations of AI in handling the emotional dimension of supervision.

A mentor from the Faculty of Medicine and Health Sciences explained:

> *"When students lose family members, struggle financially, or feel overwhelmed, they need a human being. AI cannot provide that type of support."*

A mentor from the Faculty of Law, Humanities and Social Sciences added:

> *"Research can be isolating. Students need reassurance, especially when dealing with complex ethical or philosophical questions. No AI can replace that human reassurance."*

A mentor from the *Faculty of Management and Public Administration Sciences* explained:

> *"Supervision is not just academic. Students face personal struggles, mental health issues, financial stress, and cultural pressures. AI can never replace the warmth and presence of a human mentor."*

A mentor from the *Faculty of Engineering* summarised the shared sentiment:

> *"AI might offer resources and instant answers, but when a student breaks down or feels lost, only a human can understand and respond with empathy."*

Students echoed these views strongly. A PhD student from the Faculty of Education shared:

> *"During my proposal stage, I cried in my supervisor's office because I felt like quitting. No chatbot could have supported me the way she did. When I was discouraged about my research, my mentor reassured me and helped me regain confidence. That emotional support is something AI can never provide."*

Another Master's student from the Faculty of Economics and Finance Science noted:

> *"AI cannot pick up on my tone, frustration, or stress. A human mentor immediately notices when I am not coping."*

A Masters in medicine and health science student emphasised the ethical dimension:

> *"AI can answer questions, but it cannot 'feel' with you. It doesn't understand when I'm anxious, overwhelmed, or confused. So, if AI misinterprets my emotional state or dismisses my struggles, the consequences could be harmful. Only a human mentor should guide emotional welfare."*

### 5.3.2 Sub-theme 2: Ideal features of a hybrid AI–human mentorship model

Participants across all faculties expressed interest in an integrated hybrid mentorship model where AI and humans complement each other.

A mentor from the Faculty of Law, Humanities and Social Sciences described the optimal balance:

> *"AI should handle admin tasks, reminders, early feedback, and resource sharing. This frees me to focus on higher-order feedback, idea development, and emotional support. However, the human mentor should handle intellectual engagement and wellbeing."*

A mentor from the Faculty of Medicine and Health Sciences suggested:

> *"AI could generate reading lists, track student progress, and highlight patterns in draft submissions. But conceptual discussions must remain human-led."*

A mentor from the Faculty of Engineering also suggested:

> *"AI is a useful tool for enhancing scalability and supporting large supervision cohorts. However, a machine will not replace a human. Therefore, Using AI for scalability and administrative tasks then combining a human mentor will certainly produce good results quickly."*

Students also envisioned practical uses for AI. PhD student, Faculty of Law, Humanities and Social Sciences said:

> *"I want AI to help me track deadlines, suggest article databases, and check ethics compliance forms."*

A Masters student, Faculty of Management and Public Administration Sciences suggested:

> *"AI could monitor my progress and notify my supervisor when I'm falling behind."*

However, students were firm about maintaining human contact: Another PhD student in Faculty of Education stated:

> *"AI should never be allowed to replace meetings, emotional support, or conceptual supervision. Those must remain human."*

A Masters student from the Faculty of Engineering emphasised the boundaries:

> *"AI should support, not replace. It cannot understand my personal struggles or reassure me when I am uncertain."*

One Masters postgraduate student in the Faculty of Economics and Finance Science summarised the shared view across faculties:

> *"AI should assist, not replace. It must do the routine work so that my supervisor can do the real mentoring."*

Participants across all faculties strongly converged on one central point: AI cannot replace human mentorship when it comes to emotional support, empathy, trust-building, motivation, and the understanding of personal or contextual challenges. Mentors in Education, Law, Humanities and Social Sciences, and Medicine and Health Sciences particularly emphasised the non-negotiable human role in offering emotional reassurance and contextual guidance. PhD students across all faculties echoed this, noting that the emotional strain of research requires human connection. However, divergences emerged concerning the specific roles AI should play in a sustainable mentorship framework. Mentors in the Law, Humanities and Social Sciences faculties stressed ethical concerns, including bias, privacy, and the danger of students becoming overly dependent on AI. Meanwhile, mentors from the Faculty of Engineering viewed AI as a useful tool for enhancing scalability and supporting large supervision cohorts, particularly where technical or data-driven feedback is required. The level of study also shaped responses. Master's students preferred AI for organisational tasks such as reminders and literature classification, whereas PhD students preferred AI strictly as an administrative aid, not a conceptual advisor. Departments with high supervision loads, such as Education, expressed the strongest interest in

AI-supported frameworks that reduce supervisory bottlenecks. These differences highlight that while a hybrid model is generally accepted, its design must be discipline-sensitive, ethically grounded, and aligned with student needs at different levels.

## 6. Discussion of Findings

This study aimed to design a sustainable framework that combines AI-driven effectiveness with human-centred mentorship practices, thereby allowing an optimal equilibrium of technological advancement and customised guidance. To be able to achieve the aim, the following questions were asked:

- How can Artificial Intelligence tools be integrated into academic mentorship to enhance efficiency while preserving the essential human factors of empathy, intuition, and contextual relevance?
- What are the key challenges and benefits experienced by academic mentors and mentees when engaging with hybrid AI-human mentorship models?
- How can a sustainable framework for AI-supported mentorship be developed to ensure inclusivity, scalability, and ethical practices in diverse academic environments?

The findings emphasise the necessity of preserving the human aspects of mentorship, such as empathy, trust, and emotional support, while acknowledging that AI can potentially increase efficiency, particularly in administrative tasks for mentors. Both academic mentors and postgraduate students praised the idea of a hybrid model in which AI handles routine duties while human mentors focus on providing individualised, emotionally supportive counsel. The literature consistently emphasises the historical significance of human-centred mentorship founded on empathy, intuition, and relational depth (Deane et al., 2022). Consistent with earlier literature emphasising the relational depth of mentorship (Atchley et al., 2024), the interview data show that mentees thrive when trust, empathy, and emotional presence are prioritised, reinforcing long-standing evidence on the importance of supportive academic relationships.

This reinforces the findings indicating that mentees flourish in emotionally supportive contexts where mentors serve as intellectual and personal guides (Lechuga, 2011). Furthermore, participants were cautiously optimistic about AI technologies but were clear that these tools could never replace the delicate and profoundly relational aspects of academic mentorship. There was widespread agreement that AI might enhance mentorship as a supporting tool, particularly in scheduling, feedback, and resource allocation. From both the literature (Holmes et al., 2019) and the interviews, it is clear that AI tools improve efficiency and relieve administrative burdens, especially in scheduling, automated feedback, and performance tracking. These findings align with recent research showing that AI excels in automation and cognitive support tasks but lacks the contextual judgement and emotional capacity needed for relational mentoring roles (Woo & Cho, 2025). This supports socio-technical perspectives, where technology augments human effort rather than replaces it (Ciriello et al., 2024). However, mentors and mentees caution against overreliance, which the literature warns can lead to depersonalisation and ethical issues such as bias and a lack of context-sensitivity (Binns, 2018). Furthermore, interviewees expressed concern regarding the limitations of AI, echoing literature on emotional intelligence and ethical judgement gaps in AI systems (Selwyn, 2019).

These concerns demand a cautious, ethical design that incorporates human oversight, cultural sensitivity, and feedback loops to prevent marginalisation or algorithmic harm (Mubashir et al., 2025). Additionally, the findings support the notion that a hybrid mentorship approach that combines AI efficiencies with human intuition has the potential to enhance mentorship in higher education. However, this approach must be carefully designed to ensure that AI does not overshadow the important human traits mentees value in their intellectual and emotional growth (Woo & Cho, 2025). The literature affirms the need for a hybrid mentorship model that blends AI efficiencies with human relational depth (Luckin et al., 2016; Walker et al., 2020). The interviews confirm that mentors want AI to support but not replace them, while students appreciate how AI can streamline tasks without compromising emotional connection.

## 7. Conclusions and Recommendations

The current chapter investigates the changing landscape of academic mentorship, emphasising the use of AI-driven systems alongside traditional human-centred mentoring approaches to develop a sustainable and inclusive mentorship framework suitable for higher education institutions. The study utilises a mixed-methods approach, which includes a systematic literature review and semi-structured interviews with academic mentors and postgraduate mentees. The findings reveal that AI technologies offer significant scalability, efficiency, and administrative relief benefits. However, the human element of mentorship comprises empathy, intuition, and contextual guidance, which are irreplaceable. Furthermore, the findings show that AI can effectively supplement mentorship by automating routine tasks, including feedback delivery, scheduling, and performance tracking. Conversely, mentors and mentees emphasise that effective academic connections are founded on trust, emotional support, and personalised guidance, which AI cannot fully replicate. These insights support the development of a hybrid AI-human mentorship framework based on socio-technical systems theory, ensuring that technological advancements complement rather than replace the fundamental human values of mentoring. Figure 1 below illustrates the proposed framework: sustainable AI-supported mentorship model.

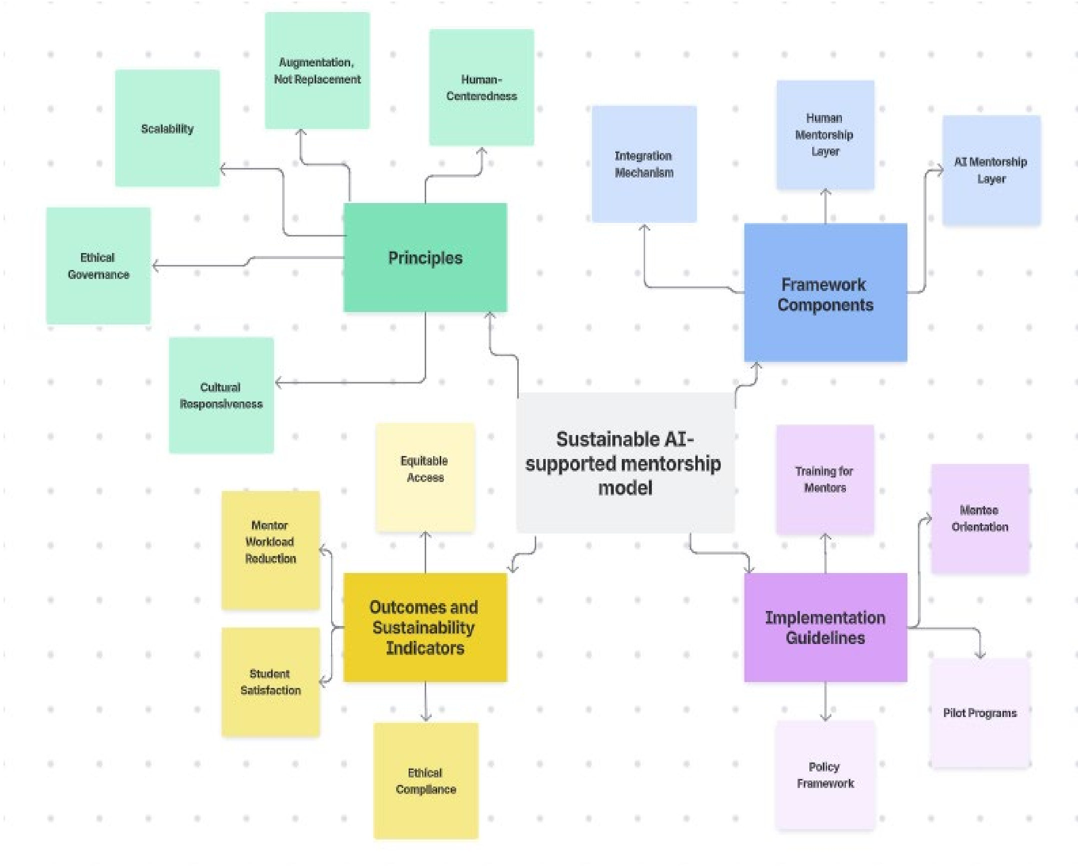



***Figure 1:*** *Proposed framework: sustainable AI-supported mentorship model*

The proposed sustainable hybrid mentorship framework incorporates socio-technical systems theory and human-centred design to ensure inclusivity, scalability, and ethical grounding. The framework is structured around three key components. The human mentorship layer provides personalised academic guidance, emotional support, contextual advice, and guidance on ethical decision-making. In contrast, the AI mentorship layer offers administrative support, resource curation, automated feedback, and performance tracking. These layers are integrated through feedback loops that allow mentees to evaluate AI interactions, explainable AI tools that clarify recommendations, and an ethical review board that oversees fairness, inclusivity, and transparency. Implementation is supported through mentor training in digital literacy and AI ethics, mentee orientation on AI use and boundaries, pilot programmes to refine processes, and institutional policies that safeguard accessibility and equity.

Expected outcomes of the framework include improved equitable access, reduced administrative burdens for mentors without compromising the quality of relationships, enhanced student satisfaction through increased engagement and trust, and consistent ethical compliance, verified through periodic audits. By combining AI-driven efficiencies with human intuition and relational depth, this framework offers a sustainable approach to mentorship that is inclusive, scalable, and ethically grounded, while preserving the personal and contextual integrity essential to effective academic supervision. Based on these, the following recommendations were made:

- The study recommends that higher education institutions develop a blended mentorship system in which AI handles the administrative and cognitive load tasks while mentors provide emotional, ethical, and contextual guidance.
- Higher education institutions should consider providing academic mentors with training on the ethical use of AI in mentorship programs, with a specific emphasis on data privacy, algorithmic bias, and the need for human presence in academic mentor-mentee partnerships.
- Educational institutions should initiate trial programs that use AI-powered mentorship tools and include robust feedback and monitoring systems to analyse impact, inclusivity, and satisfaction.
- Ongoing collaboration among mentors, mentees, and technologists is required to guarantee that mentorship innovations stay relevant, adaptable, and aligned with institutional ideals and mentees' requirements.
- Sustainable mentorship in the age of AI requires a purposeful, well-balanced, and ethical approach that takes into account both the promise of technology and the tremendous significance of human interaction. By accepting this dichotomy, institutions may improve the quality, reach, and effect of academic mentoring in an increasingly digital environment.

### 7.1 Social and Practical implications

This study carries important social implications for higher education by promoting more equitable, inclusive, and supportive mentorship experiences for postgraduate students. By integrating AI with human-centred mentorship, the framework helps reduce disparities in access to quality supervision, particularly for first-generation, rural, or historically marginalised students who often encounter inconsistent support. The model also foregrounds empathy, trust, and cultural responsiveness, ensuring that mentorship remains relational and contextually sensitive despite the introduction of technology. It strengthens students' sense of belonging, psychological safety, and overall academic confidence while advancing ethical awareness around issues such as bias, transparency, and data protection in AI-supported educational environments.

Practically, the study provides universities with a structured and scalable approach to improving mentorship systems. The framework offers clear guidance on how AI can reduce administrative burdens such as scheduling, resource curation, and early-stage feedback, allowing mentors to concentrate on higher-level academic and emotional support. At the institutional level, the model supports strategic planning by outlining training, policy development, and pilot implementation pathways that enable the safe and ethical adoption of AI tools. The framework enhances the efficiency, consistency, and quality of postgraduate mentorship while preserving the human elements essential to successful academic supervision.

## 8. Declarations

**Funding:** This research did not receive any external funding.

**Conflict of Interest**: The author declares no conflict of interest.

**Use of Artificial Intelligence:** The current work was created with the assistance of artificial intelligence technologies QuillBot Version 4.17.0 and Grammarly to assist with refining language for clarity.

## References

Ajani, O., Gamede, B., & Matiyenga, T. (2024). Leveraging artificial intelligence to enhance teaching and learning in higher education: Promoting quality education and critical engagement. *Journal of Pedagogical Sociology and Psychology, 7*(1), 54–69. https://doi.org/10.33902/JPSP.202528400

Akinwalere, S., & Ivanov, V. (2022). Artificial intelligence in higher education: Challenges and opportunities. *Border Crossing, 12*(1), 1–15. https://doi.org/10.33182/bc.v12i1.2015

Alalawi, K., Athauda, R., & Chiong, R. (2024). An extended learning analytics framework integrating machine learning and pedagogical approaches for student performance prediction and intervention. *International Journal of Artificial Intelligence in Education,* 1–49. https://doi.org/10.1007/s40593-024-00429-7

Atchley, P., Pannell, H., Wofford, K., Hopkins, M., & Atchley, R. (2024). Human and AI collaboration in the higher education environment: Opportunities and concerns. *Cognitive Research: Principles and Implications*, *9*(1), 1-11. https://doi.org/10.1186/s41235-024-00547-9

Baba, K., El Faddouli, N., & Cheimanoff, N. (2024). Mobile-optimized AI-driven personalized learning: A case study at Mohammed VI Polytechnic University. *International Journal of Interactive Mobile Technologies, 18*(4), 81–96. https://doi.org/10.3991/ijim.v18i04.46547

Bao, W. (2020). COVID-19 and online teaching in higher education: A case study of Peking University. *Human Behavior and Emerging Technologies, 2*(2), 113–115. https://doi.org/10.1002/hbe2.191

Binns, R. (2018). Fairness in machine learning: Lessons from political philosophy. *Proceedings of the 2018 Conference on Fairness, Accountability and Transparency,* 149–159. https://doi.org/10.1145/3287560.3287598

Chan, C., & Tsi, L. (2023). *The AI revolution in education: Will AI replace or assist teachers in higher education?* https://www.researchgate.net/publication/370469331_The_AI_Revolution_in_Education_Will_AI_Replace_or_Assist_Teachers_in_Higher_Education

Ciriello, R. F., Richter, A., & Mathiassen, L. (2024). Emergence of creativity in IS development teams: A socio-technical systems perspective. *International Journal of Information Management, 74,* 1–14. https://doi.org/10.1016/j.ijinfomgt.2023.102698

Coetzee, D. (2023). *Education as a conduit: Connecting individual potential to societal needs.* https://doi.org/10.13140/RG.2.2.10104.42248

Curtin, N., Malley, J., & Stewart, A. J. (2016). Mentoring the next generation of faculty: Supporting academic career aspirations among doctoral students. *Research in Higher Education, 57*(6), 714–738. https://doi.org/10.1007/s11162-015-9403-x

Deane, K. L., Boat, A. A., Haddock, S. A., Henry, K. L., Zimmerman, T. S., & Weiler, L. M. (2022). The comparative roles of mentor self-efficacy and empathy in fostering relationship quality with youth. *Journal of Youth and Adolescence, 51*(4), 805–819. https://doi.org/10.1007/s10964-022-01584-7

Donner, M. & Hummel, S. (2025). *Systematic Literature Review of AI-based Mentoring in Higher Education.* https://doi.org/10.13140/RG.2.2.26529.52320

Dwivedi, D. (2025). Emotional intelligence and artificial intelligence integration strategies for leadership excellence. *Advances in Research, 26*(1), 84–94. https://doi.org/10.9734/air/2025/v26i11235

Essop, A. (2021). *COVID-19: The "new normal." and the future of higher education.* https://www.uj.ac.za/wp-content/uploads/2021/09/covid-19-and-the-future-of-higher-education.pdf

Farrelly, T., & Baker, N. (2023). Generative artificial intelligence: Implications and considerations for higher education practice. *Education Sciences, 13*(11), 1–14. https://doi.org/10.3390/educsci13111109

Fitria, T. (2021). Artificial intelligence (AI) in education: Using AI tools for teaching and learning process. In *Prosiding Seminar Nasional & Call for Paper STIE AAS,* Surakarta, Jawa Tengah. https://www.researchgate.net/publication/357447234_Artificial_Intelligence_AI_In_Education_Using_AI_Tools_for_Teaching_and_Learning_Process

Goel, A., & Polepeddi, L. (2016). *Jill Watson: A virtual teaching assistant for online education.* Georgia Institute of Technology. https://research.gatech.edu/jill-watson

Gumede, N., & Tladi, B. (2023). Socio-technical systems: Using actor-network theory to make the social more tangible in a technical space. *South African Journal of Industrial Engineering, 34*(3), 140–155. https://doi.org/10.7166/34-3-2957

Hagler, M. (2023). Mentoring first-generation college students: Examining distinct relationship profiles based on interpersonal characteristics, support provision, and educational capital. *Journal of Community Psychology, 51*(8), 3103–3120. https://doi.org/10.1002/jcop.23003

Holmes, W., Bialik, M., & Fadel, C. (2021). *Artificial intelligence in education: Promises and implications for teaching and learning.* Center for Curriculum Redesign.

Ifraheem, S., Rasheed, M., & Siddiqui, A. (2024). Transforming education through artificial intelligence: Personalization, engagement, and predictive analytics. *Journal of Asian Development Studies, 13*(2), 250–266. https://doi.org/10.62345/jads.2024.13.2.22

Iglesias-Pradas, S., Hernández-García, Á., Chaparro-Peláez, J., & Prieto, J. (2021). Emergency remote teaching and students' academic performance in higher education during the COVID-19 pandemic: A case study. *Computers in Human Behavior, 119,* 106713. https://doi.org/10.1016/j.chb.2021.106713

Imamguluyev, R., Hasanova, P., Imanova, T., Hajizade, S., & Samadova, Z. (2024). AI-powered educational tools: Transforming learning in the digital era. *International Research Journal of Modernization in Engineering Technology and Science, 6*(12), 920–929. https://doi.org/10.56726/IRJMETS65040

Johnson, W. B., & Ridley, C. R. (2018). The elements of mentoring (3rd ed.). St. Martin's Press.

Judijanto, L. (2025). Beyond Access: Cultural, Ethical, and Infrastructural Challenges of AI in Marginalised Education Contexts. *European Journal of Contemporary Education and E-Learning, 3*(6), 83-98. https://doi.org/10.59324/ejceel.2025.3(6).07

Jyoti, J., & Sharma, P. (2015). Impact of mentoring functions on career development: Moderating role of mentoring culture and mentoring structure. *Global Business Review, 16*(4), 700–718. https://doi.org/10.1177/0972150915581110

Kaye, A. (2024). Compassion-focused mentoring: An antidote to adverse faculty-student mentoring outcomes? *OBM Integrative and Complementary Medicine, 9*(2), 1–22. https://doi.org/10.21926/obm.icm.2402024

Khumalo, H., & Ndlovu, A. K. (2024). Mentoring as a form of transformation in academia. *African Journal of Inter/Multidisciplinary Studies, 6*(S1), 1–10. https://doi.org/10.51415/ajims.v6i1.1425175

Kuznetsova-Bogdanovitsh, K., & Jyrämä, A. (2022). The role of mentorship in universities: The knowledge management framework. In *Proceedings of the 23rd European Conference on Knowledge Management (ECKM 2022).* https://papers.academic-conferences.org/index.php/eckm/article/download/470/466/2058

Lechuga, V. M. (2011). Faculty–graduate student mentoring relationships: Mentorship in a multicultural context. *Higher Education, 62*(6), 757–771. https://doi.org/10.1007/s10734-011-9416-0

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education.* Pearson Education.

Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust.* Pantheon.

Mbanjwa, S. T. (2025). Supervisory overload and its impact on postgraduate student success: A cross-sectional case study of selected institutions. *Student's Journal of Health Research Africa, 6*(6), 1-8. 3006-1059. https://doi.org/10.51168/sjhrafrica.v6i6.1743

Mohamed Badawy, D. W., Mahfouz, A., Alkahtani, A., & Melweth, A. (2023). The role of artificial intelligence technologies in enhancing education and fostering emotional

intelligence for academic success. *Migration Letters, 20*(S9), 863–874. https://doi.org/10.59670/ml.v20iS9.4868
Mubashir, M., Rana, A., & Shahnawaz, B. (2025). Human and ai collaboration in the higher education environment: Opportunities and concerns. *Contemporary Journal of Social Science Review, 3*(2), 2095-2099. https://doi.org/10.63878/cjssr.v3i2.829
Ncube, M., & Ngulube, P. (2024). Optimising data analytics to enhance postgraduate student academic achievement: A systematic review. *Education Sciences, 14*(11), 1–25. https://doi.org/10.3390/educsci14111263
Nkoala, A. S. (2024). Balanced-Integration: A dimension of supervision to support students navigating parenthood in pursuit of a PhD. *South African Journal of Higher Education, 38*(1), 189-216. https://dx.doi.org/10.20853/38-1-6270
Ntshongwana, Z. (2024). The importance of mentorship opportunities for women in academia: A systematic review. *African Journal of Inter/Multidisciplinary Studies, 6*(S1), 1–12. https://doi.org/10.51415/ajims.v6i1.143022
Nuis, W., Segers, M., & Beausaert, S. (2023). Conceptualizing mentoring in higher education: A systematic literature review. *Educational Research Review, 41,* 1–22. https://doi.org/10.1016/j.edurev.2023.100565
Odularu, O. I., & Akande, J. O. (2024). Reflection of experiences with academic supervisors, supervisees, and issues of power. *South African Journal of Higher Education, 38*(2), 254–273. https://doi.org/10.20853/38-2-5706
Owen, H. (2015). Making the most of mobility: Virtual mentoring and education practitioner professional development. *Research in Learning Technology, 23,* 1–14. https://doi.org/10.3402/rlt.v23.25566
Peiser, G., Ambrose, J., Burke, B., & Davenport, J. (2018). The role of the mentor in professional knowledge development across four professions. *International Journal of Mentoring and Coaching in Education, 7*(1), 2–18. https://doi.org/10.1108/IJMCE-07-2017-0052
Rachmad, Y. (2022). The role of artificial intelligence and human collaboration in management 5.0: A global perspective. The Academic Professor Board Protocol No 786-P07754-7385346 Feb 2022, United Nations Global Compact 137635. https://doi.org/10.17605/osf.io/vfmkr
Sachpasidi, C., Bompani, B., & Nicol, C. (2024). Intercultural encounters: Mentorship relations as spaces for critical intercultural learning in higher education institutions. *Journal of Studies in International Education, 29*(1), 64–81. https://doi.org/10.1177/10283153241275035
Saha, A., & Mondal, C. (2024). *Artificial intelligence in education: Revolutionizing learning and teaching.* Red Shine. https://www.researchgate.net/publication/383073512_ARTIFICIAL_INTELLIGENCE_IN_EDUCATION_Revolutionizing_Learning_and_Teaching

Sajja, R., Sermet, Y., Cikmaz, M., Cwiertny, D., & Demir, I. (2024). Artificial intelligence-enabled intelligent assistant for personalized and adaptive learning in higher education. *Information, 15*(10), 1–23. https://doi.org/10.3390/info15100596

Sambunjak, D., Straus, S. E., & Marušić, A. (2010). A systematic review of qualitative research on the meaning and characteristics of mentoring in academic medicine. *Journal of General Internal Medicine, 25*(1), 72–78. https://doi.org/10.1007/s11606-009-1165-8

Sarany, C., Raju, D., & Rajakumar, G. (2023). The role of a mentor in students' personal growth, academic success, and professional development. *London Journal of Research in Humanities and Social Sciences, 22*(17), 35–46. https://www.researchgate.net/publication/367090752_The_Role_of_a_Mentor_in_Students'_Personal_Growth_Academic_Success_and_Professional_Development

Selwyn, N. (2019). *Should robots replace teachers? AI and the future of education.* Polity Press.

Shen, J., DiPaola, D., Ali, S., Sap, M., Park, H. W., & Breazeal, C. (2024). Empathy toward artificial intelligence versus human experiences and the role of transparency in mental health and social support chatbot design: comparative study. *JMIR Mental Health, 11*, 1-13. https://doi.org/10.2196/62679

Singe, S., Sheldon, L., Rynkiewicz, K., Manning, C., Filep, E., Zuk, E., & Hargrave, C. (2021). Mentorship experiences of doctoral students: Understanding desired attributes of doctoral student mentors. *Internet Journal of Allied Health Sciences and Practice, 19*(1), 1–9. https://doi.org/10.46743/1540-580X/2021.1967

Singh, A. K. & Singh, T. (2025). AI-Tutors: Can AI replace human teachers. *International Journal of Statistics and Applied Mathematics, 10*(6): 145-153. https://www.doi.org/10.22271/maths.2025.v10.i6b.2068

Slimi, Z. (2023). The impact of artificial intelligence on higher education: An empirical study. *European Journal of Educational Sciences, 10*(1), 17–33. https://doi.org/10.19044/ejes.v10n1a17

Stozhko, N., Bortnik, B., Sudakova, N., & Stozhko, D. (2021). The mentoring system as a factor in the sustainable development of higher education. In *Proceedings of the Second Conference on Sustainable Development: Industrial Future of Territories (IFT 2021)* (Vol. 195). https://www.researchgate.net/publication/356658596_The_Mentoring_System_as_a_Factor_in_the_Sustainable_Development_of_Higher_Education

Straus, S., Johnson, M., Marquez, C., & Feldman, M. (2013). Characteristics of successful and failed mentoring relationships: A qualitative study across two academic health centers. *Academic Medicine, 88*(1), 82–89. https://doi.org/10.1097/ACM.0b013e31827647a0

Stravakou, P., & Lozgka, E. (2022). Postgraduate students’ views on mentoring in the educational and academic context. *European Journal of Education and Pedagogy, 3*(3), 42–51. https://doi.org/10.24018/ejedu.2022.3.3.319

Suman, D., Karthik, A., & Tenny, E. (2024). Optimizing administrative efficiency and student engagement in education: The impact of AI. *International Journal of Current Science Research and Review, 7*(10), 7792–7804. https://doi.org/10.47191/ijcsrr/V7-i10-34

Tsai, Y. S., Poquet, O., Gašević, D., Dawson, S., & Pardo, A. (2020). Complexity leadership in learning analytics: Drivers, challenges, and opportunities. *British Journal of Educational Technology, 51*(6), 1207–1222. https://doi.org/10.1111/bjet.12846

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology, 45*(3), 223–235. https://doi.org/10.1080/17439884.2020.1798995

Woo, H., & Cho, Y. Y. (2025). Human–AI collaboration: Students' changing perceptions of generative artificial intelligence and active learning strategies. *Sustainability*, 17(18), 8387. https://doi.org/10.3390/su17188387

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 1–27. https://doi.org/10.1186/s41239-019-0171-0

CHAPTER ELEVEN

# Transforming Research Supervision: Ethical and Literacy Imperatives in the Era of Generative AI

**Gardner Mwansa**[1]
**Matipa Ricky Ngandu**[2]

AFFILIATIONS

[1&2]Department of Networking and Information Technology Support, Faculty of Engineering, Built Environment and Information Technology, Walter Sisulu University, East London, South Africa.

CORRESPONDENCE

Email: rngandu@wsu.ac.za

REFERENCE

Mwansa, G., & Ngandu, M. R. (2026). Transforming research supervision: Ethical and literacy imperatives in the era of generative AI. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 191-217). ERRCD Forum. https://doi.org/10.38140/obp4-2026-11



**Abstract:** Higher education has undergone a rapid transformation in recent years, driven by the dual pressures of mitigating the long-term effects of COVID-19 and integrating generative artificial intelligence (GenAI) technologies. The pandemic exposed and exacerbated pre-existing inequalities and power imbalances within the sector, necessitating policy adaptations to address issues such as digital inequality, limited social interaction, barriers faced by student researchers in conducting face-to-face data collection, and the protection of mental health. Concurrently, GenAI has emerged as a disruptive technology that is reshaping pedagogical practices, research processes, and supervisory relationships. Although GenAI is widely promoted as a tool that can enhance teaching, research, administration, and student support, it raises critical concerns related to academic integrity, ethics, systemic bias, knowledge ownership, and uneven regulatory standards. Supervisors similarly hold divergent views regarding its usefulness and risks, a tension also reflected in inconsistent journal policies on GenAI use. Guided by the GenAI–Technological Pedagogical Content Knowledge framework (GenAI-TPACK), this study examined the ethical and literacy imperatives necessary for transforming research supervision in the era of GenAI. A systematic literature review was conducted to identify emerging GenAI literacy indicators that facilitate ethical, transparent, responsible, and informed engagement with GenAI during the research process. The review revealed significant gaps in supervisor preparedness, uneven AI literacy among research candidates, and a lack of coherent institutional guidance. The study contributes practical insights for higher education institutions seeking to balance the opportunities and challenges posed by GenAI and offers direction for developing humanising, context-sensitive guidelines for responsible integration in research supervision.

***Keywords:*** GenAI literacy, GenAI-TPACK, generative artificial intelligence, higher education, humanising pedagogy, research supervision.

## 1. Introduction

In 2022, the rapid emergence of generative artificial intelligence (GenAI) within the higher education sector, particularly following the release of tools such as ChatGPT, constituted a significant turning point in the transformation of postgraduate research (Capano et al., 2025). This phenomenon promptly generated considerable interest, as well as serious concerns regarding its potential implications. The opportunities presented by GenAI are posited as tools that can enhance higher education through personalised learning, efficient research, streamlined administration, and improved student support (Kutty et al., 2024; Noroozi et al., 2024).

The COVID-19 pandemic posed fewer opportunities and more challenges for online postgraduate supervision. The few identified opportunities included the necessity for supervisors to rethink and redesign more sustainable supervisory strategies capable of withstanding new disruptions to postgraduate research, such as those introduced by GenAI. Consequently, research candidates who pursued postgraduate studies during the pandemic experienced a heightened sense of care, guidance, and support from their supervisors. Although this support was rendered virtually, the increased level of engagement fostered favourable experiences for the candidates (Sosibo, 2024). The pandemic also provided research candidates with access to technology-based communication platforms, facilitating more frequent interactions with supervisors. Virtual communities of practice were established to enable one-on-one and group discussions. Additionally, the pandemic created an environment in which research candidates had more time to engage with research activities that would ordinarily have been consumed by work, family, and other life commitments (Chigona & Sosibo, 2024).

However, despite the opportunities introduced by GenAI and the COVID-19 pandemic, both phenomena have exposed and intensified pre-existing inequalities and power dynamics within higher education. Their impact has accelerated the necessity for policy adaptations to address digital inequality, diminished social interaction, barriers to face-to-face data collection for student researchers, and concerns related to mental health. The rethinking and redesigning of postgraduate supervision in the era of GenAI should be underpinned by humanising pedagogy. Supervisors must remain sensitive to the individual needs and backgrounds of their students to ensure that newly developed strategies are impactful and beneficial for all research candidates (Khene, 2014). This is particularly crucial in contexts where students may encounter cultural, social, and socio-economic barriers. This study aims to examine the ethical and literacy imperatives necessary for transforming research supervision in the era of GenAI within the principles of humanising pedagogy.

### 1.1 Ethical imperatives

In this era of GenAI, ethical considerations have become paramount as these technologies increasingly reshape norms and professional paradigms within the realm of postgraduate supervision. The emergence of GenAI introduces ethical dilemmas that necessitate a comprehensive understanding of its impact on ownership, creativity, bias, decision-making, and administration. Ethical imperatives must address issues of accountability, transparency, bias, and fairness as GenAI evolves and becomes more embedded in society (Dabis & Csáki, 2024).

Building on these foundational ethical concerns, ethical frameworks emphasise the importance of maintaining human oversight and accountability in the use of GenAI. This ensures that AI tools are employed responsibly and that individuals remain morally and legally accountable for AI-related outcomes. The ethical implications of utilising GenAI in social science research encompass concerns regarding misinformation, biases, privacy, and data rights (Dabis & Csáki, 2024; Saleem et al., 2024). Navigating the moral terrain of GenAI requires higher education institutions to implement policies that combine preventive measures with procedures that encourage or compel research candidates to voluntarily address the ethical concerns raised by the use of GenAI (Saleem et al., 2024).

In addition to accountability and oversight, transparency in the utilisation of GenAI is crucial. Clear communication about the role and extent of GenAI usage in research processes helps maintain trust and integrity (Dabis & Csáki, 2024; Saleem et al., 2024). Addressing algorithmic bias through the diversification of training data and ensuring fairness in GenAI applications remains an ongoing challenge. Ethical guidelines and robust protocols are also required to prevent discrimination, privacy breaches, and other forms of dehumanisation. According to Runcan et al. (2025) and Silva-Atencio (2025), GenAI raises ethical dilemmas related to governance, bias, transparency, and fairness, underscoring the need for responsible AI frameworks and proactive ethical guidance in research.

Beyond considerations of governance and transparency, GenAI has the potential to enhance data collection and analysis, offering alternative approaches to traditional research methodologies, which highlights the necessity for comprehensive digital and AI literacy training in research supervision (Farina & Stevenson, 2024). Globally, higher education institutions are developing comprehensive ethical guidelines to govern the use of GenAI. These guidelines are essential to align the integration of GenAI with academic integrity and social responsibility (Farina & Stevenson, 2024; Montezuma & Chong, 2024). Ethical reflection and continuous moral education are vital to assist research candidates and supervisors in navigating the complexities of GenAI usage in research (Montezuma & Chong, 2024; Saleem et al., 2024).

### 1.2 AI Literacy imperatives

The integration of GenAI into research processes presents distinctive opportunities and challenges for both research candidates and supervisors, thereby necessitating enhanced AI literacy among these groups. GenAI possesses the potential to exacerbate the digital divide and adversely affect access to research resources (Radojičić & Vukmirović, 2025). AI literacy can play a crucial role in mitigating this challenge by equipping research candidates and supervisors with the skills and knowledge required to engage with GenAI effectively, ethically, and responsibly (Ruiz et al., 2024).

To address these emerging literacy gaps, there exists a significant need for training programmes aimed at enhancing AI literacy among research candidates and supervisors. This encompasses understanding how to generate effective prompts and critically evaluate AI outputs (Takaffoli et al., 2024). Higher education institutions are encouraged to adapt their research processes and pedagogical approaches to better prepare future researchers for the AI era. Petrenko (2024) emphasises the importance of developing skills such as scenario-based thinking and uncertainty management.

Beyond individual training and skill development, fostering interdisciplinary interactions and diverse viewpoints is essential for the ethical and effective use of GenAI. Collaboration among technologists, educators, and social scientists can assist in addressing the multifaceted ethical issues posed by GenAI. The deployment of GenAI introduces the potential for biases and other ethical challenges that may impact the cognitive engagement of participants or researchers (Saleem et al., 2024). This highlights the necessity for careful ethical consideration and reinforces the value of essential human capabilities within the research process.

At an institutional level, establishments must evaluate their preparedness for GenAI adoption, taking into account factors such as data readiness, ethical safeguards, and leadership support. The successful integration of AI necessitates a balance between technological innovation and ethical considerations (Marcinkevage & Kumar, 2025). GenAI has the potential to transform postgraduate research and alleviate the workload of research candidates and supervisors, underscoring the need for future research to explore effective pedagogical approaches and the long-term impacts of GenAI integration (Farina & Stevenson, 2024). The use of GenAI systems in academia raises ethical queries concerning accountability, transparency, bias, and fairness as these systems advance in sophistication, highlighting both the advantages and challenges of maintaining moral principles in research supervision (Dabis & Csáki, 2024). GenAI also raises issues regarding privacy, consent, and responsible data management, necessitating further investigation into effective security measures and responsible data practices (Montezuma & Chong, 2024). The challenges associated with integrating GenAI into research supervision include inaccuracies, usability, privacy concerns, and the indispensable role of human review, underscoring the need for guidance and strategies for the appropriate and ethical integration of GenAI in literature reviews (Takaffoli et al., 2024).

Therefore, when considered collectively, transforming research supervision in the era of GenAI necessitates a dual focus on ethical imperatives and literacy imperatives. By promoting a culture of ethical responsibility and enhancing AI literacy, higher education institutions can harness the potential of GenAI while mitigating risks associated with human factors through humanising pedagogies. While existing studies provide valuable insights into the ethical and literacy imperatives for research supervision in the era of GenAI, they do not directly address the challenges of integrating GenAI into research supervision from an ethical, literacy, and humanising perspective.

### 1.3 Humanising pedagogy in postgraduate supervision in an era of GenAI

The primary objective of humanising pedagogy within a digital context is to cultivate learner dignity, establish meaningful interactions, and promote learner empowerment, all within a complex digital landscape (Mehta & Aguilera, 2020). The significance of digital pedagogy lies in its ability to effectively address the urban-rural digital literacy divide through comprehensive skills development initiatives. Humanising pedagogy in supervision requires supervisors to remain sensitive to the individual needs and backgrounds of their students, particularly in contexts where students may face cultural and social barriers (Khene, 2014). While mentoring is an effective supervision strategy, it is essential to recognise the power dynamics involved. Effective supervision should not obscure the significant role of power but rather address it transparently (Manathunga, 2007). Technology plays a significant role in modern research, and the emergence of GenAI is reshaping traditional boundaries between supervisors and research candidates. GenAI can offer personalised learning experiences by generating tailored learning materials and providing immediate feedback, which can help address the unique needs of each postgraduate student (Kong & Yang, 2024; Rajak et al., 2024). GenAI can support the diverse learning styles of research candidates, as its technologies can create inclusive learning environments by offering personalised support and adaptive learning platforms, which are particularly beneficial for all learning styles (Velazquez-Solis et al., 2025).

The challenges and considerations of GenAI use in research processes include ethical use and academic integrity. The widespread availability of GenAI tools raises concerns about academic misconduct. Higher education institutions need to develop comprehensive policies that balance the benefits of GenAI with the need to uphold established academic standards (Kruger-Roux & Alberts, 2024; Strachan et al., 2024). Tensions in power dynamics in postgraduate supervision require a rethink, as GenAI may be perceived as a superior source of information and guidance. Supervisors must become increasingly conscious of the power dynamics in their relationships with research candidates (Manathunga, 2007).

Integrating GenAI into postgraduate supervision can enhance the learning experience by providing personalised and immediate support. However, literature has shown that it is crucial to maintain a humanising pedagogy that is sensitive to the individual needs of research candidates and aware of the power dynamics involved. By balancing the benefits of GenAI with ethical considerations and continuous professional development, supervisors can create a supportive and inclusive learning environment for their research candidates.

## 2. Conceptual Framework

This study adopts a conceptual framework grounded in the Technological Pedagogical Content Knowledge (TPACK) theory originally proposed by Mishra and Koehler (2006). TPACK extends pedagogical content knowledge by integrating technological knowledge to explain how educators meaningfully combine technology, pedagogy, and content in educational practice. While TPACK has been widely applied in teaching and teacher education contexts, recent studies demonstrate its adaptability for examining emerging digital technologies, including artificial intelligence, within higher education environments (Celik, 2023; Valtonen et al., 2017).

Building on this theoretical foundation, the present study advances an adapted GenAI-TPACK conceptual framework tailored to postgraduate research supervision. This framework does not claim to constitute a new theory; rather, it synthesises established theoretical constructs and contemporary literature to explain how generative artificial intelligence influences supervisory practice under specific ethical, pedagogical, and institutional conditions.

The framework, illustrated in Figure 1, comprises three interrelated knowledge dimensions: GenAI literacy, humanising pedagogy, and postgraduate supervision, all situated within an outer contextual layer of ethical institutional governance.

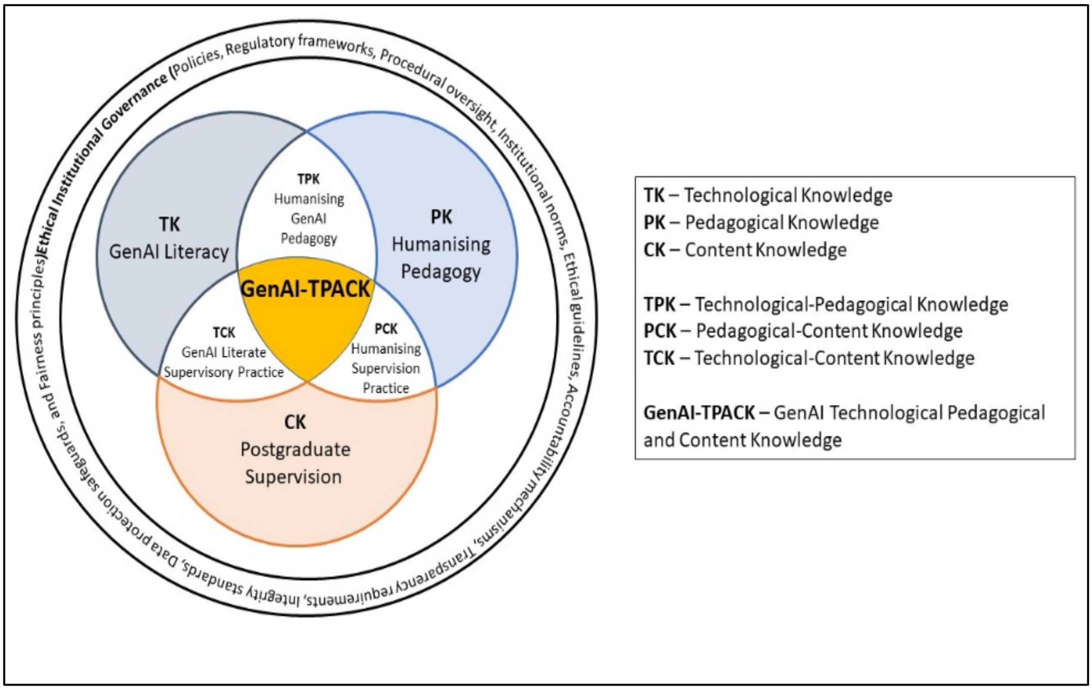



***Figure 1:*** *Conceptual Framework for Ethical and Human-Centred GenAI-Mediated Postgraduate Supervision*

GenAI literacy encompasses the competencies required by supervisors and postgraduate candidates to engage critically and responsibly with GenAI tools. Prior research conceptualises AI literacy as extending beyond technical proficiency to include evaluative judgement, ethical awareness, transparency, and the ability to recognise bias and limitations in AI-generated outputs (Celik, 2023; Ng et al., 2021). In the context of postgraduate supervision, insufficient GenAI literacy may result in an inappropriate reliance on automated outputs, compromised academic integrity, and inconsistent supervisory practices.

Humanising pedagogy serves as the pedagogical foundation of the framework. Drawing on scholarship that emphasises relationality, care, dialogue, and learner agency in higher education (Freire, 2021; Kim et al., 2023), this dimension highlights the importance of preserving trust, dignity, and reflexive engagement within AI-mediated supervision. Recent studies caution that AI-supported educational practices must complement, rather than replace, the socio-cultural and dialogic elements that underpin effective supervision (Kukulska-Hulme et al., 2022).

Postgraduate supervision constitutes the content domain of the framework and encompasses the disciplinary, methodological, and developmental practices involved in guiding research candidates through the research process. Literature on doctoral and postgraduate supervision highlights the centrality of feedback, scholarly socialisation, identity formation, and epistemic development, all of which may be reshaped by GenAI-enabled tools (Boyd & Harding, 2025; Lee, 2008; Wisker et al., 2010). This dimension interacts dynamically with GenAI literacy, as supervisors increasingly integrate digital tools into supervision while remaining accountable for academic standards and learning outcomes.

Surrounding these three dimensions is the outer layer of ethical institutional governance, which provides the structural and regulatory context for GenAI adoption. Existing research consistently identifies gaps in institutional policies, disclosure requirements, accountability mechanisms, and ethical guidance for AI use in higher education (Azevedo et al., 2025; Floridi et al., 2018; UNESCO et al., 2023). Ethical institutional governance encompasses policies, regulatory frameworks, procedural oversight, institutional norms, ethical guidelines, transparency requirements, and data protection safeguards. This layer ensures that the responsibility for ethical GenAI use does not rest solely with individual supervisors or students but is supported by coherent institutional structures.

As illustrated in Figure 1 above, these four dimensions operate as an integrated model:

- GenAI literacy enables critical and responsible engagement with AI tools;
- Humanising pedagogy ensures that supervision remains relational, ethical and student-centred;
- Postgraduate supervision provides disciplinary and developmental grounding; and
- Ethical institutional governance establishes the conditions for accountable and equitable GenAI integration.

Together, these elements form a holistic conceptual framework that explains how GenAI shapes postgraduate supervisory practice and identifies the institutional and ethical conditions required for its responsible adoption. The framework provides a structured foundation for analysing supervisory practices, institutional readiness, literacy disparities and the relational consequences of GenAI integration in postgraduate research supervision.

## 3. Problem Statement

Recent shifts in higher education have been influenced by rapid digital evolution, which has the potential to enhance access to learning, increase flexibility, and foster more personalised educational experiences that enrich both teaching and research practices (Zou et al., 2025). These developments, including heightened reliance on digital tools and remote learning environments, have generated new expectations regarding the conduct of research and research supervision. Existing studies indicate that digital transformation can facilitate improved learning environments that promote independent learning, adaptive engagement, and data-informed decision-making. However, the extent to which these advantages are realised depends on institutional readiness, professional competencies, and strategic vision within higher education institutions (Ajani, 2024; Nazyrova et al., 2025). Within this evolving digital context, GenAI has emerged as a significant disruptor in higher education (Francis et al., 2025; García-López & Trujillo-Liñán, 2025; He, 2025; Jin et al., 2025). Its integration has further redefined research practices, with digital tools increasingly regarded not merely as complements to traditional approaches, but as integral components of research support and postgraduate supervision. While GenAI presents considerable potential for enhancing research efficiency and allowing researchers to concentrate on more complex facets of their work through personalised support and administrative automation, its rapid and uneven adoption has introduced a range of tensions that warrant critical examination. In particular, the pace of GenAI development, combined with predominantly small-scale and experimental utilisation within supervisor-research candidate relationships across both the global north and south (Petrenko, 2024), has resulted in fragmented regulatory responses, limited preparedness among supervisors and research candidates, and missed opportunities for research excellence. This situation highlights a pronounced gap in understanding the ethical and GenAI literacy requirements necessary to facilitate responsible and effective GenAI integration in postgraduate supervision. Accordingly, this study provides new insights into the ethical and GenAI literacy imperatives essential for supervisors and research candidates to navigate the evolving landscape of postgraduate supervision in the context of widespread GenAI adoption.

### 3.1 Research questions

This study is guided by the following three research questions, which integrate the ethical, literacy, governance, and pedagogical considerations emerging from the literature:

- Q1. What ethical challenges arise from the integration of GenAI into postgraduate research supervision, particularly regarding academic integrity and mentorship dynamics?
- Q2. What GenAI literacy competencies and institutional governance mechanisms are required to ensure responsible, ethical, and consistent use of GenAI by research supervisors and postgraduate students?
- Q3. What strategies can align the adoption of GenAI in postgraduate research supervision with humanising pedagogy while preserving the integrity of knowledge creation?

## 4. Materials and Methods

A systematic literature review (SLR) was conducted for this study, employing the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methodology. The PRISMA 2020 statement was applied (Page et al., 2021). PRISMA facilitates the documentation of transparent accounts of the conducted review, with its utilisation widely endorsed and adopted within the academic community. The rigour, comprehensiveness, and reproducibility inherent in PRISMA render it particularly suitable for this type of investigation. The researchers identified and classified the research as a systematic review report addressing the ethical and literacy imperatives in the current era of Generative AI and its impact on transforming postgraduate research supervision. To ensure a rigorous and transparent review process, clearly defined inclusion and exclusion criteria were established for the selection of literature pertinent to this study. These criteria guided the identification, screening, and refinement of sources, ensuring that only relevant, credible, and recent publications contributed to the analysis. Table 1 summarises the specific parameters utilised to determine the eligibility of studies for inclusion and the rationale for exclusion.

***Table 1:*** *Inclusion and exclusion criteria*

| Inclusion | Exclusion |
|---|---|
| 1. Search keywords | 1. Unrelatedness to search keywords |
| 2. Articles written in English | 2. Articles in other languages except English |
| 3. Academically peer-reviewed articles, conferences, books, book chapters, reviews | 3. Not peer-reviewed articles |
| 4. Recency (2015-2025) | 4. Recency (>10 years) |
| 5. Duplicates | 5. Duplicates |
| 6. Article or title relatedness to the phenomenon under investigation | 6. Discipline of focus |
| 7. Relatedness to the problem statement, research question, and objective | 7. Outside the domain of postgraduate supervision |

The researchers conducted searches across Google Scholar, Scopus, the Education Resources Information Centre (ERIC), and the Web of Science databases. The rationale for selecting these databases was to ensure a comprehensive and rigorous search, utilising sources that grant access to a substantial body of relevant literature and are specifically designed to index academic research in education and related fields. Scopus and Web of Science are renowned for their extensive coverage of high-quality peer-reviewed publications and robust citation indexing capabilities (Martín-Martín et al., 2021). ERIC is widely acknowledged as a fundamental database for educational research due to its emphasis on scholarly and policy-related outputs (Fitzgerald et al., 2025). Google Scholar was included to expand the search, as it encompasses grey literature and additional academic outputs that may not be represented in conventional subscription databases (Halevi et al., 2017). Table 2 presents a list of the databases utilised, the search strings applied, and the resulting number of initial sources (112).

Table 2: Search strings

| # | Database | Search term | Number of sources |
|---|---|---|---|
| 1 | Google Scholar | ("Generative AI" OR "GenAI" OR "Artificial Intelligence") AND ("research supervision" OR "doctoral supervision" OR "graduate mentorship") AND ("ethics" OR "academic integrity" OR "responsible AI" OR "bias") AND ("AI literacy" OR "digital literacy" OR "supervisor training") AND ("humanizing pedagogy" OR "humanising pedagogy" OR "social justice" OR "humanising praxis" OR "humanizing praxis") | 89 |
| 2 | Scopus | ("Generative AI" OR "GenAI" OR "Artificial Intelligence") AND ("research supervision" OR "doctoral supervision" OR "graduate mentorship") | 11 |
| 3 | ERIC | ("Generative AI" OR "GenAI" OR "Artificial Intelligence") AND ("research supervision" OR "doctoral supervision" OR "graduate mentorship") | 8 |
| 4 | Web of Science | (ALL=("Generative AI" OR "GenAI" OR "Artificial gence")) AND ALL=("research supervision" OR "doctoral vision" OR "graduate mentorship") | 4 |
| | | **Total:** | **112** |

Table 2 presents the results from each database search. The researchers subsequently identified articles published between 2015 and 2025, resulting in the exclusion of 18 articles. Further screening was conducted on the remaining 94 articles, leading to the exclusion of 6 duplicates. Additionally, 53 articles were excluded for failing to meet the specified criteria, as their abstracts or titles were not related to the phenomenon under investigation. A further 25 articles were excluded due to their lack of relevance to the problem statement and research questions. Ultimately, only 10 articles remained for final analysis based on the inclusion criteria, with none sourced from Web of Science. Although only ten studies fulfilled the inclusion criteria, this number is deemed appropriate given the emerging and conceptually specialised nature of research on GenAI in postgraduate research supervision. Systematic review methodologists emphasise that small evidence bases are common and methodologically acceptable in developing or narrowly defined research areas, particularly when rigorous eligibility criteria are employed to ensure conceptual relevance and analytical depth (Brunton et al., 2012, 2017; Petticrew & Roberts, 2006). Importantly, the objective of a systematic literature review (SLR) is not to maximise the number of included studies but to ensure transparency, coherence, and alignment between the research purpose and the selection process. This review does not aim to estimate prevalence, effectiveness, or generalisable outcomes; rather, it seeks to synthesise ethical, literacy, and pedagogical insights that enhance understanding of GenAI-mediated postgraduate supervision. In conceptually oriented reviews, the adequacy of the evidence base is assessed by conceptual saturation and analytical coherence rather than numerical thresholds alone (Snyder, 2019; Xiao & Watson, 2019). Consequently, the inclusion of ten studies reflects the point at which no substantively new conceptual insights emerged, indicating sufficient saturation to address the three research questions. The limited number of included studies is therefore not a methodological weakness but a notable finding in itself, highlighting both the novelty of the field and the necessity for further empirical research.

In this study, articles addressing artificial intelligence (AI) in education at a general level were deliberately excluded to preserve conceptual integrity and maintain close alignment with the research questions focused specifically on ethical and GenAI literacy imperatives in postgraduate supervision. In accordance with PRISMA 2020, the rigour of the review is demonstrated through methodological transparency, reproducibility, and the defensible application of inclusion and exclusion criteria, rather than the attainment of a predetermined sample size (Page et al., 2021). Adhering to these principles, the final sample of ten studies represents the most relevant, methodologically appropriate, and conceptually aligned evidence currently available for the phenomenon under investigation (Albright, 2011; Booth et al., 2016). Figure 2 below illustrates the flowchart representing the PRISMA process followed (Page et al., 2021).

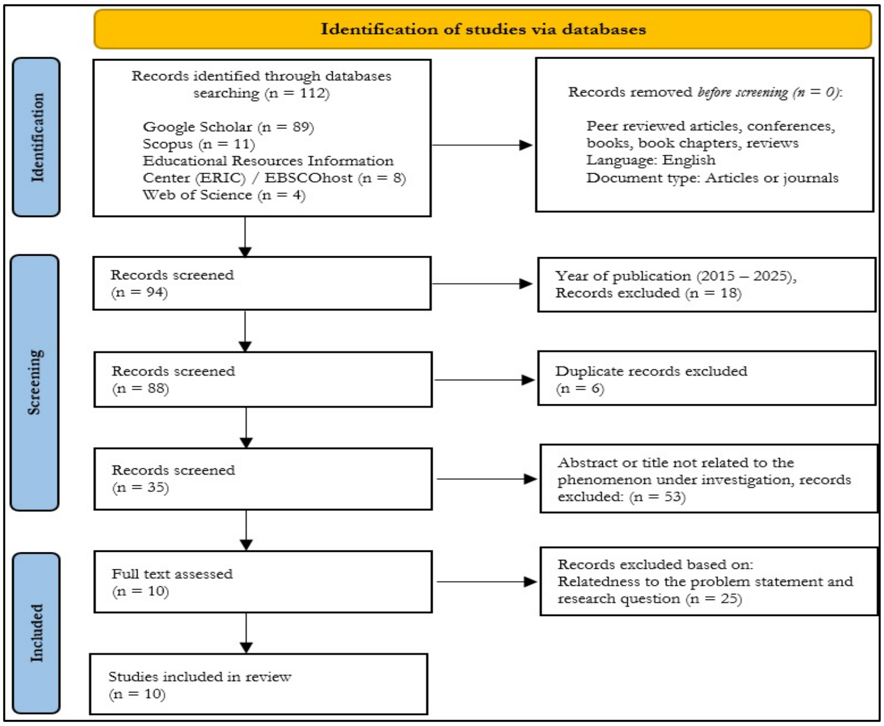



***Figure 2:*** *PRISMA flow diagram*

## 5. Presentation of Results

This section presents the findings of the SLR based on the ten studies that met the inclusion criteria. As detailed in the Materials and Methods section, this modest evidence base reflects the emerging and highly specialised nature of scholarship on ethical and GenAI literacy imperatives in postgraduate supervision. Methodological guidance confirms that, in conceptually narrow or developing fields, smaller study pools remain acceptable when derived through rigorous screening and clearly defined inclusion criteria (Brunton et al., 2012, 2017; Petticrew and Roberts, 2006; Snyder, 2019; Xiao & Watson, 2019).

The findings presented in this section derive from a SLR undertaken to examine the ethical and GenAI literacy imperatives shaping postgraduate research supervision in higher education, viewed through a humanising pedagogy lens. The review was guided by the overarching research question: What ethical and GenAI literacy imperatives are necessary for transforming research supervision in higher education through a humanising pedagogy in the era of GenAI?

Although the study employs three research questions, the thematic analysis identifies four themes. This is consistent with systematic review conventions in which themes arise inductively from the evidence base. In this case, the second research question, integrating GenAI literacy and governance, produced two distinct clusters: one centred on competency development and the other on institutional regulation. Because the literature treats these as conceptually distinct domains, they are presented separately to ensure analytic clarity. These four dimensions informed the subsequent thematic synthesis and continue to serve as the organising structure for presenting the results.

A total of ten validated sources were included in the analysis, summarised in Tables 3 and 4. Each source was examined against the four analytical dimensions that correspond directly to the study's research questions, as indicated earlier. Table 2 illustrates the frequency with which each of the ten studies addressed these dimensions, showing that GenAI literacy (n = 8) and ethical challenges (n = 7) were the most prominent areas of focus, followed by humanising pedagogy (n = 6) and policy and governance (n = 4). This distribution informed the thematic organisation of the findings. The results are therefore structured around four themes, each reflecting one of the research questions and the corresponding analytical dimension. These themes synthesise patterns across the reviewed literature and highlight areas of convergence and divergence relevant to GenAI-mediated postgraduate supervision.

*Table 3: Descriptive summary of sources*

| N0: | Author(s) | Title | Year | Publisher | Publisher Country | Type of Article |
|---|---|---|---|---|---|---|
| 1 | Bartoletti, I. | Chapter 3 – AI in education | 2022 | Routledge | UK | Book chapter |
| 2 | Smuha, N.A. | Chapter 5 – Pitfalls and pathways for Trustworthy AI | 2022 | Routledge | UK | Book chapter |
| 3 | Kizilcec, R. F. & Lee, H. | Chapter 7 – Algorithmic fairness in education | 2022 | Routledge | UK | Book chapter |
| 4 | du Boulay, B. | Chapter 9 – The overlapping ethical imperatives of human teachers and their Artificially Intelligent assistants | 2022 | Routledge | UK | Book chapter |

| | | | | | | |
|---|---|---|---|---|---|---|
| 5 | Jassim, O.A., Mahmoud, M.A. & Sharifuddin, M. | A multi-agent framework for research supervision management | 2015 | Springer Verlag | Germany | Journal article |
| 6 | Waring, P. | Artificial intelligence and graduate employability: What should we teach Generation AI? | 2024 | Kaplan Higher Education | Singapore | Journal article |
| 7 | Tauginienė, L. | Ethics and Integrity in Research, Publishing, and Educational Leadership: Introduction | 2024 | Springer Nature | Germany | Journal article |
| 8 | Fu, Yao & Weng, Zhenjie | Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centred AI practices | 2024 | Elsevier | Netherlands | Journal article |
| 9 | Wright, A. | Postgraduate Supervision in a ChatGPT World: What's Next? | 2024 | Universidad Politecnica de Valencia | Spain | Conference paper |
| 10 | Cowling, M., Crawford, J., Allen, K. & Wehmeyer, M. | Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research supervision. | 2023 | Australasian Journal of Educational Technology (AJET) | Australia | Journal article |

Table 3 compiles 10 scholarly sources (2015–2025) that contribute to the discourse on Artificial Intelligence in education, algorithmic ethics, research supervision, and graduate employability. These sources are categorised by type, publisher details, and origin to support systematic literature synthesis. There is a balanced mix of empirical research, conceptual chapters, and emerging discourse contributions, comprising 5 journal articles, 4 book chapters, and 1 conference paper. The sources reflect a broad international representation, particularly from Europe. The topical focus of the sources ranges from Artificial Intelligence ethics and fairness in education, such as algorithmic bias and human-AI interaction, to postgraduate and graduate teaching supervision, with an emphasis on evolving practices in AI-enhanced academic contexts; digital pedagogy; research integrity and mentorship, with a focus on ethics, leadership, and professional development in graduate education. Most works are from 2022–2024, showing a very recent and relevant research base on Artificial Intelligence and postgraduate pedagogy in higher education. This collection of sources forms a solid foundation for a systematic review on the evolving role of Artificial Intelligence in higher education, especially regarding ethics, employability, and postgraduate supervision.

Table 4 further deepens the analysis by mapping each of the ten validated studies to the three research questions (Q1–Q3) using keyword alignment and conceptual relevance. This table illustrates how each article contributes differently across the ethical, literacy, policy, and humanising pedagogy dimensions. For example, studies by Bartoletti (2022), Smuha (2022), Kizilcec and Lee (2022), and du Boulay (2022) show strong alignment with Q1 and Q2, reflecting their emphasis on ethical concerns and GenAI literacy. Similarly, works by Cowling et al. (2023)

and Wright (2024) align with Q3 due to their emphasis on relational, supervisory, and pedagogical implications. This structured mapping provides the basis for identifying patterns across the literature and informs the thematic synthesis presented later in this section.

***Table 4:*** *Summary of Sources*

| N0: | Author(s) | Keywords Matched | Q 1 | Q 2 | Q 3 | Summary |
|---|---|---|---|---|---|---|
| 1 | Bartoletti, I. | AI, ethics, education, policy | ✓ | ✓ | ✓ | Primary theme: AI in education<br><br>Focus / contribution: Ethical implications of AI integration in education |
| 2 | Smuha, N.A. | trustworthy AI, ethics, integrity, fairness | ✓ | ✓ | ✓ | Primary theme: Trustworthy AI<br><br>Focus / contribution: Critically evaluates the risks and challenges to building trustworthy AI in education |
| 3 | Kizilcec, R. F. & Lee, H. | algorithmic bias, equity, fairness, pedagogy | ✓ | ✓ | ✓ | Primary theme: Algorithmic fairness<br><br>Focus / contribution: Explores fairness and bias in educational algorithms |
| 4 | du Boulay, B. | AI ethics, academic integrity, transparency | ✓ | ✓ | | Primary theme: Ethics of Human-AI Roles<br><br>Focus / contribution: Analyses the ethical imperatives of teachers vs AI assistants |
| 5 | Jassim, O.A., Mahmoud, M.A. & Sharifuddin, M. | AI supervision, agent systems, governance, doctoral training | | ✓ | ✓ | Primary theme: AI in Supervision<br><br>Focus / contribution: Proposes a multi-agent framework to support research supervision processes |
| 6 | Waring, P. | AI literacy, employability, higher education | | ✓ | | Primary theme: AI and employability<br><br>Focus / contribution: Examines AI's influence on graduate employability and curriculum design |
| 7 | Tauginienė, L. | academic integrity, leadership, ethics | ✓ | ✓ | | Primary theme: Research ethics and integrity<br><br>Focus / contribution: Introduces ethical leadership in research and publishing |
| 8 | Fu, Y. & Weng, Z. | responsible AI, educational ethics, human-centred design | ✓ | | ✓ | Primary theme: Responsible human-centred AI<br><br>Focus / contribution: Presents a systematic review on frameworks for ethical, responsible AI in education |
| 9 | Wright, A. | GenAI, research supervision, integrity | ✓ | ✓ | ✓ | Primary theme: Postgraduate supervision in AI context<br><br>Focus / contribution: Explores how AI (e.g., ChatGPT) is reshaping postgraduate supervision |
| 10 | Cowling, M., Crawford, J., Allen, K. & | ChatGPT, research supervision, humanising pedagogy | | ✓ | ✓ | Primary theme: AI and leadership in supervision<br><br>Focus / contribution: Analyses leadership strategies for integrating AI in undergraduate and postgraduate research supervision |

| | | | | | | |
|---|---|---|---|---|---|---|
| | Wehmeyer, M. | | | | | |
| **Distribution of supporting articles:** | | | 7 | 8 | 7 | |

The presentation of findings is structured by themes. The four themes outlined below synthesise patterns, tensions, and convergences across the included studies. Collectively, they illuminate the ethical, pedagogical, literacy, and governance imperatives necessary for embedding GenAI within postgraduate supervision in a manner consistent with a humanising pedagogy. The themes integrate both qualitative insights and keyword trend analyses to highlight patterns, gaps, and areas of transformative potential in research supervision. The discussion incorporates theoretical grounding from the GenAI-TPACK framework and situates the findings within the context of humanising pedagogy for postgraduate supervision in the era of GenAI.

### 5.1 Theme 1: Ethical challenges in GenAI-mediated postgraduate supervision

Seven studies illustrate the complex ethical implications of integrating GenAI into postgraduate supervision. Across this body of work, the central concern is the potential erosion of academic integrity and uncertainty surrounding authorship when GenAI-generated content becomes embedded in research processes. Bartoletti (2022) contends that undisclosed or poorly regulated use of GenAI risks compromising originality and diminishing the scholarly value of research outputs. This concern is reinforced by Smuha (2022), who notes that many institutional contexts still treat AI tools as neutral, thereby overlooking the socio-political dynamics and rights-based considerations crucial to trustworthy AI practices.

Bias and fairness also emerge as critical challenges. Kizilcec and Lee (2022) demonstrate that algorithmic systems can reproduce existing societal inequities, which has implications for supervisory evaluation, feedback, and decision-making processes. Their work aligns with du Boulay (2022), who critiques the implicit value systems encoded in educational AI tools and stresses the necessity for proactive ethical reflection rather than reactive compliance. These ethical risks, including algorithmic opacity and compromised integrity, highlight the need for clearer institutional support and transparent accountability structures.

Leadership and governance perspectives further reinforce this theme. Tauginienė (2024) emphasises that fragmented ethical cultures in universities leave supervisors without guidance, resulting in inconsistent decisions regarding authorship, disclosure, and fairness. Fu and Weng (2024) also reveal governance gaps, arguing that responsible AI use must be grounded in human-centred ethical principles rather than performative policy statements. Wright (2024) provides empirical evidence that supervisors already experience uncertainty and stress when navigating these ambiguities without institutional clarity. Collectively, these studies underscore that the ethical challenges of GenAI are multifaceted, affecting integrity, transparency, fairness, and relational trust throughout the supervision process.

### 5.2 Theme 2: GenAI literacy competencies for supervisors and postgraduate students

Eight studies emphasise the necessity for comprehensive GenAI literacy among supervisors and postgraduate students. The literature consistently indicates that GenAI literacy extends beyond basic operational familiarity, encompassing ethical reasoning, evaluative judgement, and critical engagement with AI-generated outputs. Bartoletti (2022) identifies a significant preparedness gap in higher education, noting that the adoption of AI is accelerating faster than institutions can provide adequate training. This observation aligns with Smuha (2022), who argues that trustworthy AI necessitates scenario-based ethical training rather than superficial compliance with guidelines.

Algorithmic bias and fairness remain central considerations within the discourse on literacy. Kizilcec and Lee (2022) demonstrate that, without the ability to detect and question biased AI outputs, supervisors risk perpetuating inequities in assessment and feedback. Their findings complement du Boulay's (2022) assertion that AI literacy and ethical literacy must be cultivated concurrently, as users must critically engage with GenAI rather than treating it as an infallible technological assistant. This dual competency is essential for maintaining rigorous academic standards.

Employability-driven perspectives also contribute to this theme. Waring (2024) warns that graduates lacking GenAI literacy may encounter difficulties in AI-driven labour markets, thereby underscoring the importance for universities to embed AI reasoning, prompt engineering, and critical digital skills into curricula. At the institutional level, Tauginienė (2024) observes that literacy gaps create unequal supervisory experiences, reinforcing the need for structured AI capacity-building initiatives. Wright (2024) adds that supervisors frequently feel unsupported when evaluating AI-influenced student work, further emphasising the lack of training frameworks. Cowling et al. (2023) illustrate the applied dimension of literacy by demonstrating that supervisors must understand prompt design, validation, and oversight to integrate GenAI responsibly. Collectively, these studies depict GenAI literacy as a multidimensional competency essential for ethical, consistent, and informed supervisory practices.

### 5.3 Theme 3: Institutional policies and governance frameworks for GenAI Use

Across four studies, institutional governance emerges as a critical yet underdeveloped component of the integration of GenAI in postgraduate supervision. These works collectively emphasise that the adoption of GenAI cannot be left to individual discretion and requires structured, institution-wide policies to ensure ethical and consistent practice. For instance, Jassim et al. (2015) highlight the necessity for systematised supervisory processes through their multi-agent model, while Fu and Weng (2024) demonstrate that higher education institutions often lack governance mechanisms capable of keeping pace with rapid advancements in GenAI. Similarly, Tauginienė (2024) and Brunton et al. (2012) stress the importance of institutional leadership, coherence, and ethical oversight to support supervisors navigating AI-mediated environments.

A recurring message across these studies is that governance gaps heighten risks related to bias, academic integrity, uneven supervision practices, and misalignment between institutional expectations and supervisory realities. Fu and Weng (2024) caution that without clear guidelines, the burden of ethical decision-making shifts to individual supervisors, creating inconsistencies that may compromise fairness and accountability. Complementing this, Tauginienė (2024) argues that robust integrity safeguards, transparent disclosure expectations, and institutional support structures are essential to prevent fragmented or reactive responses to the challenges posed by GenAI.

Overall, these studies indicate that institutional policy and governance are not peripheral considerations but foundational elements of the ethical implementation of GenAI in postgraduate supervision. In the absence of coherent frameworks, both students and supervisors are left to manage complex ethical questions independently, thereby increasing the likelihood of inconsistent practices and compromised academic standards (Fu & Weng, 2024; Jassim et al., 2015; Tauginienė, 2024).

### 5.4 Theme 4: Humanising pedagogy and relational dynamics in AI-supported supervision

Six studies explore how GenAI can be integrated into postgraduate supervision in ways that preserve dignity, relational trust, and student agency—core principles of a humanising pedagogy. The literature consistently highlights that GenAI offers potential benefits but also risks undermining the interpersonal aspects of supervision if not embedded within relationally grounded pedagogical practices. Bartoletti (2022) emphasises that AI tools must supplement rather than replace supervisor–student dialogue, cautioning that misuse may dehumanise learning processes. This view is complemented by Smuha (2022), who argues that ethical AI governance must foreground human rights, well-being, and socio-cultural sensitivity.

Fairness and inclusivity also shape this theme. Kizilcec and Lee (2022) demonstrate that algorithmic systems may marginalise certain groups if fairness and transparency are not intentionally prioritised. Their call for participatory design processes aligns with Fu and Weng (2024), who report that most current educational AI systems lack human-centred design principles, potentially reinforcing power imbalances within supervisory relationships. Both studies highlight the need for pedagogy that ensures AI enhances rather than suppresses student expression and epistemic agency.

Practice-oriented insights strengthen the pedagogical dimension of this theme. Wright (2024) illustrates that supervisors often feel pressure to adopt GenAI tools without adequate guidance, risking tensions between institutional expectations and humanising supervisory commitments. She proposes co-designed supervision agreements to maintain transparency and relational accountability. Similarly, Cowling et al. (2023) demonstrate that GenAI can support formative feedback and student development when integrated within leadership models that prioritise psychological safety, collaboration, and relational care. Together, these studies underscore that humanising pedagogy remains essential for ensuring that GenAI enhances rather than diminishes the ethical, relational, and developmental foundations of postgraduate supervision.

## 6. Discussion of Findings

The systematic review demonstrates that the integration of GenAI into postgraduate research supervision presents both significant opportunities and complex ethical, pedagogical, and governance challenges. The ten validated articles collectively suggest that while GenAI can enhance personalisation, productivity, and formative feedback processes, its implementation occurs within institutional contexts that often lack clear policy direction, consistent literacy training, and ethically grounded supervisory practices. When interpreted through the broader literature and the conceptual arguments established in the Introduction, several cross-cutting insights emerge.

Interpreting the findings through the GenAI-TPACK lens highlights how responsible GenAI-mediated supervision depends on the interplay between technological proficiency, pedagogical sensitivity, and deep disciplinary knowledge. This framework is particularly relevant in a rapidly transforming academic environment where supervisors must reconsider traditional roles, expectations, and relational practices in light of GenAI's capabilities and constraints. Moreover, the findings align with humanising pedagogy, which foregrounds relational trust, inclusivity, socio-cultural awareness, and student voice; elements that remain central to postgraduate supervision even as technological mediation increases.

Ethical concerns emerge as a foundational theme. Articles by Bartoletti (2022), Smuha (2022), du Boulay (2022), and Tauginienė (2024) converge on the argument that issues of academic integrity, attribution, and algorithmic bias require urgent attention in supervisory contexts. Despite this, many institutions still lack ethical clarity regarding how GenAI outputs should be acknowledged, evaluated, or incorporated into thesis writing, data analysis, and feedback cycles. These findings echo Dabis and Csáki (2024) and Saleem et al. (2024), who emphasise the need for stronger institutional safeguards and more robust oversight mechanisms. The absence of clear guidance contributes to differing interpretations among supervisors and students, thereby influencing relational trust and expectations.

The importance of GenAI literacy among supervisors and postgraduate researchers is another dominant theme. Eight validated studies, including those by Kizilcec and Lee (2022), Waring (2024), and Cowling et al. (2023), highlight that literacy extends beyond technical skill to include ethical reasoning, critical prompt design, bias recognition, and reflective integration of AI-generated outputs. The findings align with calls from Farina and Stevenson (2024) and Petrenko (2024), who emphasise scenario-based training and institutional readiness. The contribution from Wright (2024) underscores the vulnerability supervisors experience when left to navigate the implications of AI without centralised support structures or institutional backing.

Despite the increasing adoption of GenAI in academic discourse, policy frameworks remain limited. Only four validated articles, specifically those by Jassim et al. (2015), Fu and Weng (2024), and Tauginienė (2024), provide a substantive discussion of governance structures. This observation is consistent with insights from Montezuma and Chong (2024), who note a persistent gap between policy intentions and regulatory enforcement. The absence of cohesive institutional policy not only inhibits responsible implementation but also raises equity concerns regarding differential access, uneven expectations, and inconsistent supervisory practices across departments and faculties.

A critical theme arising from both validated and referenced literature is the importance of aligning GenAI use with humanising pedagogical values. Six of the selected articles, particularly those by Wright (2024), Cowling et al. (2023), and Smuha (2022), argue that AI should not erode the relational, inclusive, and context-sensitive elements of postgraduate supervision. Rather than viewing AI as a replacement for supervisory engagement, a humanising framework perceives it as a collaborator in fostering student agency, dignity, and voice. This notion is reinforced by foundational works introduced by Mehta and Aguilera (2020) and Manathunga (2007), who assert that any shift in supervision models must remain grounded in care, mutual respect, and ethical transparency.

While excluded or contextually aligned studies were not analysed in depth, they provide a useful interpretative context. For instance, Velazquez-Solis et al. (2025) and Kruger-Roux and Alberts (2024) expand on how inclusive GenAI systems can support diverse learning needs, particularly for students from underrepresented or resource-constrained backgrounds. These findings align with Khene's (2014) advocacy for context-sensitive pedagogy that is culturally and socially responsive. Meanwhile, the literature from the COVID-19 era, as discussed by Sosibo (2024) and Chigona and Sosibo (2024), contextualises the shift towards virtual supervision models and examines how these frameworks have set the stage for contemporary AI-mediated supervision dynamics.

The discussion surrounding power dynamics introduced in the theoretical framing of GenAI and Technological Pedagogical Content Knowledge (TPACK) remains underexplored in most empirical studies. The potential for GenAI to exacerbate asymmetries between supervisors and research candidates—especially when only one party has access to or control over AI tools—raises ethical concerns that intersect with issues of consent, authorship, and student autonomy. In summary, the convergence between the systematic literature review (SLR) findings and the broader literature suggests that the integration of GenAI into postgraduate supervision is not merely a technical upgrade but rather represents a pedagogical and ethical shift. As institutions move towards scaling GenAI adoption, they must anchor these transitions in clearly articulated ethics, widespread literacy training, inclusive governance, and a recommitment to humanising pedagogy.

## 7. Conclusion

This systematic review examined the ethical and GenAI literacy imperatives shaping postgraduate research supervision in higher education, interpreted through a humanising pedagogy framework. The findings demonstrate that the integration of GenAI presents significant opportunities for enhancing feedback, productivity, and supervisory support while simultaneously introducing complex ethical, relational, and governance challenges that institutions have not yet adequately addressed.

Across the reviewed literature, four key patterns emerged. First, ethical concerns, including academic integrity, authorship, algorithmic bias, and transparency, remain insufficiently regulated, leaving supervisors and students without clear guidance on the responsible use of GenAI. Second, GenAI literacy has become an essential competency for both supervisors and postgraduate students. However, the development of this literacy currently depends on individual initiative rather than coordinated institutional support, resulting in uneven levels of understanding and inconsistent supervision practices. Third, the review reveals a persistent policy and governance vacuum. Despite the rapid adoption of GenAI tools, few institutions have developed coherent regulatory frameworks that outline expectations, responsibilities, or acceptable boundaries for AI-assisted academic work. Finally, the findings underscore the need to integrate GenAI use within a humanising pedagogical approach. Without intentional attention to relational trust, inclusivity, and student agency, GenAI risks reinforcing existing power hierarchies and diminishing the developmental, dialogic character of postgraduate supervision.

Overall, the evidence indicates that GenAI does not merely introduce new technological capabilities; it reshapes the ethical, pedagogical, and relational foundations of supervision. Transforming postgraduate supervision in the GenAI era, therefore, requires institutional readiness, supervisor support, and pedagogical sensitivity grounded in ethical responsibility and human-centred educational values.

### 7.1 Recommendations

The findings of this review highlight several ethical, pedagogical, and governance considerations that must be addressed to support the responsible and human-centred integration of GenAI into postgraduate research supervision. Drawing from the thematic analysis and the preceding conclusion, this section outlines key recommendations for higher education institutions, supervisors, and policymakers. These recommendations are intended to strengthen institutional preparedness, enhance GenAI literacy, safeguard ethical academic practice, and ensure that technological adoption aligns with the principles of humanising pedagogy.

Higher education institutions should develop coherent and comprehensive policies that clearly articulate expectations for the ethical use of GenAI in postgraduate supervision. These policies should address disclosure procedures, authorship attribution, acceptable uses of GenAI tools in academic work, and mechanisms for fairness auditing. Establishing consistent guidelines across faculties will reduce reliance on individual interpretation and minimise the risk of unequal supervisory practices. Institutions should also implement governance structures that support ethical decision-making, provide clarity regarding roles and responsibilities, and ensure that supervisors and students engage with GenAI within a regulated and transparent framework.

There is a critical need for structured GenAI literacy programmes that equip supervisors and postgraduate students with the competencies required for responsible engagement with AI technologies. Such programmes should extend beyond basic technical skills to include ethical reasoning, critical evaluation of AI outputs, awareness of algorithmic bias, and proficiency in prompt design. Embedding GenAI literacy into postgraduate training and supervisor development initiatives will promote consistent and informed practice across academic departments. Ongoing professional development is essential, given the rapid evolution of GenAI tools and their growing influence on research practices.

Ethical engagement with GenAI in postgraduate supervision requires transparent communication between supervisors and students regarding expectations, boundaries, and acceptable forms of AI assistance. Institutions should encourage supervisors and students to co-create supervision agreements that explicitly address GenAI use and align with principles of academic integrity. Supervisors should adopt practices that integrate fairness, accountability, and human rights considerations into their decision-making processes, ensuring that GenAI enhances rather than undermines the ethical standards and relational foundations of postgraduate research.

The adoption of GenAI tools should be guided by the values of humanising pedagogy, which emphasises relational trust, student agency, socio-cultural sensitivity, and mutual respect. Supervisors should be supported to integrate GenAI in ways that augment, rather than replace, the interpersonal and developmental aspects of postgraduate supervision. This includes ensuring that students retain their voice and autonomy, and that AI tools are introduced in ways that recognise diverse learning needs and safeguard student wellbeing. Intentional alignment between GenAI use and human-centred pedagogical principles will help prevent the reinforcement of power imbalances or the erosion of meaningful supervisory relationships.

## 7.2 Future research directions

Further research is needed to deepen our understanding of how GenAI shapes postgraduate supervision across different institutional contexts. Empirical studies examining how supervisors and students negotiate authorship, trust, and accountability in AI-mediated academic work would provide valuable insights. Comparative research across diverse and resource-constrained settings could illuminate the equity-related implications of GenAI adoption. Longitudinal studies exploring how GenAI influences supervisory identity, power dynamics, and epistemic autonomy over time would also contribute significantly to the evolving discourse on AI in higher education.

### 7.3 Limitations of the study

This review has several limitations that should be acknowledged. First, although 112 records were initially identified, only ten studies met the inclusion criteria. This small number reflects both the emerging nature of research on GenAI ethics and literacy within postgraduate supervision and the narrow conceptual focus of the review. Second, the study restricted its scope to peer-reviewed literature published between 2015 and 2025, potentially excluding relevant earlier work or recent grey literature. Third, the search was limited to four major academic databases, which, although appropriate for the field, may have narrowed the breadth of the available evidence. Finally, the included studies varied in methodological design and disciplinary context, which limits the generalisability of the findings. These limitations indicate the need for further empirical and comparative research to deepen understanding of GenAI's role in postgraduate supervision.

## 8. Declarations

**Funding:** This research did not receive any external funding.
**Conflicts of Interest:** The author(s) declare no conflict of interest.
**Use of Artificial Intelligence:** The current work was created with the assistance of artificial intelligence technologies, such as Grammarly, to assist with refining language for clarity, as confirmed by the author(s).

## References

Ajani, O. A. (2024). Exploring digital transformation and future trends in higher education development across African nations. *Journal of Pedagogical Sociology and Psychology*. https://doi.org/10.33902/JPSP.202427374

Albright, D. L. (2011). Book review: Cooper, H. *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: SAGE. https://doi.org/10.1177/1049731509348800

Azevedo, L., Robles, P., Best, E., & Mallinson, D. J. (2025). Institutional policies on artificial intelligence in higher education: Frameworks and best practices for faculty. *New Directions for Adult and Continuing Education*, *10*(2), 1-23. https://doi.org/10.1002/ace.70013

Bartoletti, I. (2022). AI in education. In W. Holmes & K. Porayska-Pomsta (Eds.), *The ethics of artificial intelligence in education* (1st ed.). Routledge.

Booth, A., Sutton, A., & Papaioannou, D. (2016). Taking a systematic approach to your literature review. *Systematic Approaches to Successful Literature Review*, 9-35.

Boyd, P., & Harding, D. (2025). Generative AI: Reconfiguring supervision and doctoral research. *Buildings & Cities*, *6*(1), 1-12. https://doi.org/10.5334/bc.560

Brunton, G., Stansfield, C., & Thomas, J. (2012). Chapter 6 - Finding relevant studies. In *An introduction to systematic reviews*.

Capano, G., He, A. J., & McMinn, S. (2025). Riding the tide of generative artificial intelligence in higher education policy: An Asian perspective. *Journal of Asian Public Policy*, 1–15. https://doi.org/10.1080/17516234.2025.2450571

Celik, I. (2023). Towards intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, *138*, 107468. https://doi.org/10.1016/j.chb.2022.107468

Chigona, A., & Sosibo, L. (2024). Research and postgraduate supervision during the Coronavirus-19 pandemic: Lessons learned. *South African Journal of Higher Education*, *38*(1). https://doi.org/10.20853/38-1-6281

Cowling, M., Crawford, J., Allen, K. A., & Wehmeyer, M. (2023). Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research supervision. *Australasian Journal of Educational Technology*, *39*(4). https://doi.org/10.14742/ajet.8598

Dabis, A., & Csáki, C. (2024). AI and ethics: Investigating the first policy responses of higher education institutions to the challenge of generative AI. *Humanities and Social Sciences Communications*, *11*(1), 1006. https://doi.org/10.1057/s41599-024-03526-z

du Boulay, B. (2022). The overlapping ethical imperatives of human teachers and their artificially intelligent assistants. In W. Holmes & K. Porayska-Pomsta (Eds.), *The ethics of artificial intelligence in education* (1st ed.). Routledge.

Farina, A., & Stevenson, C. N. (2024). *Ethical navigations* (pp. 63–87). https://doi.org/10.4018/979-8-3693-1565-1.ch005

Fitzgerald, S. R., Weaver, K. D., & Droog, A. (2025). Selecting a specialised education database for literature reviews and evidence synthesis projects. *Research Synthesis Methods*, *16*(1), 30–41. https://doi.org/10.1017/rsm.2024.11

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schäfer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, *28*(4). https://doi.org/10.1007/s11023-018-9482-5

Francis, N. J., Jones, S., & Smith, D. P. (2025). Generative AI in higher education: Balancing innovation and integrity. *British Journal of Biomedical Science*, *81*. https://doi.org/10.3389/bjbs.2024.14048

Freire, P. (2021). *Pedagogy of hope: Reliving pedagogy of the oppressed. https://doi.org/10.1080/00131911.2020.1766207*

*Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. Computers and Education: Artificial Intelligence, 7, 100306. https://doi.org/10.1016/j.caeai.2024.100306*

*García-López, I. M., & Trujillo-Liñán, L. (2025). Ethical and regulatory challenges of generative AI in education: A systematic review. Frontiers in Education, 10. https://doi.org/10.3389/feduc.2025.1565938*

*Halevi, G., Moed, H., & Bar-Ilan, J. (2017). Suitability of Google Scholar as a source of scientific information and as a source of data for scientific evaluation: Review of the literature. Journal of Informetrics, 11(3), 823–834. https://doi.org/10.1016/j.joi.2017.06.005*

He, D. (2025). The Impact of Generative Artificial Intelligence on Higher Education. *Proceedings of the 2nd International Conference on Intelligent Education and Computer Technology*, 324–329. https://doi.org/10.1145/3764206.3764256

Jassim, O. A., Mahmoud, M. A., & Sharifuddin, M. (2015). A multi-agent framework for research supervision management. *Advances in Intelligent Systems and Computing*, *373*. https://doi.org/10.1007/978-3-319-19638-1_15

Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, *8*, 100348. https://doi.org/10.1016/j.caeai.2024.100348

Khene, C. (2014). Supporting a humanizing pedagogy in the supervision relationship and process: A reflection in a developing country. *International Journal of Doctoral Studies*, *9*, 073–083. https://doi.org/10.28945/2027

Kim, H. J., Kong, Y., Hernandez, C., & Soban, M. (2023). Student emotions and engagement: Enacting humanizing pedagogy in higher education. *IAFOR Journal of Education*, *11*(3), 185–206. https://doi.org/10.22492/ije.11.3.09

Kizilcec, R. F., & Lee, H. (2022). Algorithmic fairness in education. In W. Holmes & K. Porayska-Pomsta (Eds.), *The ethics of artificial intelligence in education* (1st ed.). Routledge.

Kong, S.-C., & Yang, Y. (2024). A human-centered learning and teaching framework using generative artificial intelligence for self-regulated learning development through domain knowledge learning in K–12 settings. *IEEE Transactions on Learning Technologies*, *17*, 1562–1573. https://doi.org/10.1109/TLT.2024.3392830

Kruger-Roux, H., & Alberts, R. (2024). *Generative artificial intelligence policy for academic literacy in South African higher education* (pp. 1–22). https://doi.org/10.4018/979-8-3693-1054-0.ch001

Kukulska-Hulme, A., Bossu, C., Charitonos, K., Tim, C., Rebecca, F., Elizabeth, F., Mark, G., Guitert, M., Christothea, H., Maina, M., Prietp-Blázquez, J., Bart, R., Sangrà, A., Sargent, J., Scanlon, E., & Denise, W. (2022). Innovating pedagogy 2022. In *The Open University* (Issue July).

Kutty, S., Chugh, R., Perera, P., Neupane, A., Jha, M., Li, L., Gunathilake, W., & Perera, N. (2024). Generative AI in higher education: Perspectives of students, educators and administrators. *Journal of Applied Learning & Teaching*, *7*(2). https://doi.org/10.37074/jalt.2024.7.2.27

Here is the text formatted in APA 7th edition style:


Lee, A. (2008). How are doctoral students supervised? Concepts of doctoral research supervision. *Studies in Higher Education, 33*(3). https://doi.org/10.1080/03075070802049202

Manathunga, C. (2007). Supervision as mentoring: The role of power and boundary crossing. *Studies in Continuing Education, 29*(2), 207–221. https://doi.org/10.1080/01580370701424650

Marcinkevage, C., & Kumar, A. (2025). Generative AI in higher education constituent relationship management (CRM): Opportunities, challenges, and implementation strategies. *Computers, 14*(3), 101. https://doi.org/10.3390/computers14030101

Martín-Martín, A., Thelwall, M., Orduna-Malea, E., & Delgado López-Cózar, E. (2021). Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: A multidisciplinary comparison of coverage via citations. *Scientometrics, 126*(1), 871–906. https://doi.org/10.1007/s11192-020-03690-4

Mehta, R., & Aguilera, E. (2020). A critical approach to humanising pedagogies in online teaching and learning. *The International Journal of Information and Learning Technology, 37*(3), 109–120. https://doi.org/10.1108/IJILT-10-2019-0099

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record: The Voice of Scholarship in Education, 108*(6), 1017–1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Montezuma, J. R. M., & Chong, M. (2024). Generative artificial intelligence impact on education and industry: An ethical dimension. *2024 IEEE World Engineering Education Conference (EDUNINE)*, 1–3. https://doi.org/10.1109/EDUNINE60625.2024.10500608

Nazyrova, A., Miłosz, M., Bekmanova, G., Omarbekova, A., Aimicheva, G., & Kadyr, Y. (2025). The Digital Transformation of Higher Education in the Context of an AI-Driven Future. *Sustainability*, *17*(22), 9927. https://doi.org/10.3390/su17229927

Ng, D. T. K., Leung, J. K. L., Chu, K. W. S., & Qiao, M. S. (2021). AI literacy: Definition, teaching, evaluation and ethical issues. *Proceedings of the Association for Information Science and Technology, 58*(1), 504–509. https://doi.org/10.1002/pra2.487

Noroozi, O., Soleimani, S., Farrokhnia, M., & Banihashem, S. K. (2024). Generative AI in education: Pedagogical, theoretical, and methodological perspectives. *International Journal of Technology in Education, 7*(3), 373–385. https://doi.org/10.46328/ijte.845

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, n71. https://doi.org/10.1136/bmj.n71

Petrenko, A. (2024). The role of generative artificial intelligence (GAI) in scientific research. *System Research and Information Technologies, 3*, 133–147. https://doi.org/10.20535/SRIT.2308-8893.2024.3.08

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide.* Oxford: Blackwell.

Radojičić, S., & Vukmirović, D. (2025). Exploring the impact of generative AI on digital inclusion: A case study of the e-government divide. *AI, 6*(12), 303. https://doi.org/10.3390/ai6120303

Rajak, L., Chauhan, S., & Bara, S. (2024). Transforming English pedagogy with artificial intelligence: En route to enhanced language learning. In *Artificial intelligence: A multidisciplinary approach towards teaching and learning* (pp. 216–241). Bentham Science Publishers. https://doi.org/10.2174/9789815305180124010013

Ruiz, P., Mills, K., Lee, K., Coenraad, M., Fusco, J., Roschelle, J., & Weisgrau, J. (2024). *AI literacy: A framework to understand, evaluate, and use emerging technology.* https://doi.org/10.51388/20.500.12265/218

Runcan, R., Hațegan, V., Toderici, O., Croitoru, G., Gavrila-Ardelean, M., Cuc, L. D., Rad, D., Costin, A., & Dughi, T. (2025). Ethical AI in social sciences research: Are we gatekeepers or revolutionaries? *Societies*, *15*(3), 62. https://doi.org/10.3390/soc15030062

Saleem, M., Chawla, C., & Ather, D. (2024). *Ethical considerations in the deployment of generative AI systems in academia* (pp. 245–256). https://doi.org/10.1007/978-981-97-8460-8_12

Silva-Atencio, G. (2025). The challenges and opportunities for ethics in generative artificial intelligence in the digital age. *DYNA*, *92*(236), 26–35. https://doi.org/10.15446/dyna.v92n236.117144

Smuha, N. A. (2022). Pitfalls and pathways for trustworthy AI. In W. Holmes & K. Porayska-Pomsta (Eds.), *The ethics of artificial intelligence in education* (1st ed.). Routledge.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Sosibo, L. (2024). Key factors contributing to postgraduate students' success: Making a case for coronavirus-19 and beyond. *South African Journal of Higher Education*, *38*(1). https://doi.org/10.20853/38-1-6263

Strachan, R., Oguna, C., & Oruche, U. (2024). The postgraduate student perspective on academic misconduct in the era of essay mills and generative AI: A case study from Northeast England. *2024 IEEE Global Engineering Education Conference (EDUCON)*, 1–7. https://doi.org/10.1109/EDUCON60312.2024.10578821

Takaffoli, M., Li, S., & Mäkelä, V. (2024). Generative AI in user experience design and research: How do UX practitioners, teams, and companies use GenAI in industry? *Designing Interactive Systems Conference*, 1579–1593. https://doi.org/10.1145/3643834.3660720

Tauginienė, L. (2024). Ethics and integrity in research, publishing, and educational leadership: Introduction. In *Springer International Handbooks of Education: Vol. Part F2304*. https://doi.org/10.1007/978-3-031-54144-5_138

UNESCO, Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. In *Guidance for generative AI in education and research*. https://doi.org/10.54675/ewzm9535

Valtonen, T., Sointu, E., Kukkonen, J., Kontkanen, S., Lambert, M. C., & Mäkitalo-Siegl, K. (2017). TPACK updated to measure pre-service teachers' twenty-first century skills. *Australasian Journal of Educational Technology*, *33*(3). https://doi.org/10.14742/ajet.3518

Velazquez-Solis, P. E., González Correa, M. E., Martinez, M. A., Arroyo, J. G., & Marquez, M. Y. (2025). Designing teaching strategies using artificial intelligence for neurodivergent students in higher education. In *Exploring psychology, social innovation and advanced applications of machine learning* (pp. 191–208). IGI Global. https://doi.org/10.4018/979-8-3693-6910-4.ch010

Waring, P. (2024). Artificial intelligence and graduate employability: What should we teach Generation AI? *Journal of Applied Learning & Teaching*, *7*(1), 22–25. https://doi.org/10.37074/jalt.2024.7.1.42

Wisker, G., Morris, C., Cheng, M., Masika, R., Warnes, M., Trafford, V., Robinson, G., & Lilly, J. (2010). *Doctoral learning journeys final report.*

Wright, A. (2024). Postgraduate supervision in a ChatGPT world: What's next? *10th International Conference on Higher Education Advances (HEAd'24)*, 1–8. https://doi.org/10.4995/HEAd24.2024.17244

Xiao, Y., & Watson, M. (2019). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, *39*(1), 93–112. https://doi.org/10.1177/0739456X17723971

Zou, Y., Kuek, F., Feng, W., & Cheng, X. (2025). Digital learning in the 21st century: Trends, challenges, and innovations in technology integration. *Frontiers in Education*, *10*. https://doi.org/10.3389/feduc.2025.1562391

CHAPTER TWELVE

# Assessing Research Integrity in the Age of AI: A Longitudinal Analysis Using an AI Misuse Impact Index

**Tichaona Chikore**[1]
**Farai Nyabadza**[2]

AFFILIATIONS
[1&2]Department of Mathematics and Applied Mathematics, Faculty of Science, University of Johannesburg, Johannesburg, South Africa.

CORRESPONDENCE
Email: tichaona.chikore@gmail.com

REFERENCE
Chikore, T., & Nyabadza, F. (2026). Assessing research integrity in the Age of AI: A longitudinal analysis using an AI misuse impact index. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 218–244). ERRCD Forum. https://doi.org/10.38140/obp4-2026-12



**Abstract:** The increasing adoption of artificial intelligence (AI) in academic research has reshaped scholarly practices while introducing complex ethical risks, particularly concerning research integrity and academic misconduct. This study proposes a comprehensive quantitative and empirical framework, adapted from the Cobb-Douglas production function, to model how the misuse of AI contributes to systemic quality degradation, using retractions as a proxy for integrity breaches. By leveraging longitudinal publication and retraction data from Retraction Watch and Scopus, we construct an AI misuse impact index to track the relationship between research output and integrity risks over time. Time series lag analysis reveals that retraction rates most strongly correlate with prior publication volumes at a one-year lag, indicating the rapid manifestation of AI-driven misconduct. To identify critical intervention points, we apply piecewise linear modelling to detect thresholds where retraction rates accelerate disproportionately relative to publication growth. A plagiarism tolerance threshold is established, beyond which research quality deteriorates unsustainably. Additionally, we introduce a probabilistic damage model, quantifying the risk of systemic integrity failure as AI adoption expands. Results highlight a pronounced post-2009 rise in AI-related integrity risks, with a sharp inflection in 2023 when misconduct indicators exceeded acceptable tolerance levels, signalling a system-wide ethical crisis. The study further proposes a dynamic, data-driven method for calibrating institutional plagiarism thresholds in alignment with evolving integrity risks and patterns of AI adoption. This model enables proactive monitoring and policy adjustments, linking integrity governance directly to empirical risk indicators. The findings underscore the urgent need for adaptive, transparent AI oversight frameworks within academia, ensuring that AI complements rather than undermines the ethical and intellectual foundations of research. Future research should extend this work by integrating discipline-specific AI use patterns and developing real-time academic integrity monitoring systems.

***Keywords:*** Academic integrity, AI-Mediated supervision, AI policy, artificial intelligence, ethical considerations, impact index, plagiarism tolerance, responsible AI use.

## 1. Introduction

The increasing utilisation of artificial intelligence (AI) in postgraduate supervision is reshaping academic mentorship, presenting new opportunities to enhance research efficiency and streamline administrative tasks, particularly in guiding students through their complex research journeys (Ali, 2020; Altmäe et al., 2023). AI-driven tools are already transforming various aspects of academic work, such as automating literature reviews, managing administrative duties, and providing personalised, data-driven feedback to students (Kamalov et al., 2023; Ali et al., 2024). These innovations hold significant promise, but their integration into postgraduate supervision—an inherently personalised mentorship process—raises unique challenges. Postgraduate supervision extends beyond mere knowledge transfer; it fosters the development of critical research skills, ethical reasoning, and professional growth, elements that AI may struggle to replicate or support effectively (Zawacki-Richter et al., 2019; Mauti, 2025). Despite AI's efficiency, concerns are growing regarding its impact on traditional mentoring, which emphasises human interaction, ethical guidance, and individualised development, particularly in the context of postgraduate education (Nguyen & Vuong, 2024; Zhai et al., 2024).

AI tools offer clear benefits for supporting postgraduate research; however, concerns regarding their potential to undermine academic integrity, encourage over-reliance on technology, and depersonalise supervision remain underexplored (Crawford et al., 2024; Ahmad et al., 2023). The increasing integration of AI into research processes may diminish ethical reasoning and weaken the traditional mentorship bond. Additionally, the rise in article retractions since the COVID-19 pandemic, coinciding with the emergence of tools such as ChatGPT, raises pertinent questions about AI's role in academic misconduct (Nguyen & Vuong, 2024; Al-Jahwari & Yousif, 2025).

Despite the compelling qualitative evidence linking the rapid adoption of Generative AI to rising academic misconduct, the research community faces an urgent, quantifiable crisis that remains unaddressed. The most significant gap in the current literature is the critical lack of a clear, empirical, and quantitative framework to systematically measure how AI misuse translates into systemic degradation of research quality across the enterprise. Current studies acknowledge the ethical risks but fail to model the quantitative relationship between key research inputs (for example, human effort, institutional resources, and AI utilisation) and measurable outcomes of integrity failure (for example, retractions) over time (Acosta-Enriquez et al., 2025; Papagiannidis et al., 2025).

This deficiency leaves academic institutions effectively blind to the escalating risk of systemic integrity failure. Without a rigorous quantitative model, it is impossible to identify the critical inflection points, specifically the 'plagiarism tolerance thresholds', where the rate of integrity breaches begins to accelerate disproportionately, thereby exceeding institutional capacity for effective quality control. Consequently, existing governance policies remain largely reactive, failing to provide the predictive mechanisms necessary for timely, evidence-based intervention. This leads to the central, unmet need that this study addresses: the necessity for a framework to predict and quantitatively model the risk of integrity collapse in the AI-mediated research environment.

To overcome this methodological gap and provide the necessary predictive framework, we utilise a two-stage approach. This includes an analytical model adapted from the foundational

Cobb-Douglas production function (Cobb & Douglas, 1928), as applied in related research (Baulk, 2024), coupled with piecewise linear modelling to empirically detect these critical quantitative inflexion points. To explore temporal dynamics, we apply lagged correlation analysis to determine whether retractions follow publication spikes with a delay, thereby identifying optimal intervention timing. Using piecewise linear regression (Yang et al., 2016), we aim to detect thresholds where retractions accelerate disproportionately relative to output, thereby aiding in the definition of critical risk points. We also propose a plagiarism tolerance threshold to signal when misconduct exceeds acceptable limits and test this concept through simulations comparing capped and uncapped misuse scenarios. The model examines institutional enforcement patterns and adapts tolerance levels dynamically. Finally, we develop a probabilistic model to estimate the likelihood of integrity collapse due to cumulative AI misuse, integrating damage thresholds and failure probabilities to inform early institutional interventions.

This study aims to construct a quantitative integrity risk index for artificial intelligence (AI) misuse in research by developing an AI Misuse Impact Index that integrates publication and retraction data to quantify the relationship between the growth of AI-driven research output and integrity risks. The index will serve as an empirical tool for tracking vulnerabilities in research integrity as AI adoption expands. Using piecewise linear modelling and time series lag analysis, the study seeks to identify critical thresholds and tipping points where retraction rates accelerate disproportionately in relation to publication growth, signalling systemic misconduct. Furthermore, the research will establish empirically grounded plagiarism tolerance thresholds and determine critical inflection points that indicate the onset of widespread ethical degradation within academic systems. To enhance analytical precision, a probabilistic damage model will be introduced to quantify the risk of systemic research misconduct under varying scenarios of AI adoption, capturing the dynamic interplay between research output, AI misuse, and retraction patterns as a predictive tool for monitoring emerging academic integrity crises. Finally, the study proposes a dynamic, data-driven integrity governance framework that enables institutions to adapt plagiarism tolerance thresholds and integrity policies in real time, linking empirical indicators to institutional decision-making for proactive intervention and the sustained safeguarding of research integrity in AI-augmented academic environments.

We identify key points in postgraduate supervision where the use of AI should be regulated, utilising patterns of academic misconduct predicted by machine learning models. The study evaluates AI's impact on students' research skills, ethics, and development, proposing targeted interventions to mitigate risks. It offers a measurable framework for balancing AI efficiency with ethical oversight and supports policy recommendations such as ethical AI guidelines, adaptive plagiarism thresholds, and real-time integrity monitoring to assist universities in maintaining academic standards while adopting new technologies.

## 2. Materials and Methods

Our analysis follows a two-stage methodological design. This approach is fundamentally based on the Cobb-Douglas production function (Cobb & Douglas, 1928), which quantitatively models how the combined factors of AI, human effort, and institutional resources impact research quality within AI-mediated supervision. First, a general model defines the relationship between inputs and research output. Second, the model is extended to account for variations in AI adoption across individuals and institutions. This approach captures both direct and systemic effects of AI use, allowing for empirical validation and informing policy decisions.

### 2.1 The general case

We propose a model where research quality (Q) depends on study resources (R), combined effort (E), and AI use (A). Resources include infrastructure such as libraries and software, while effort reflects contributions from both students and supervisors. AI acts as a multiplier, enhancing how resources and effort affect research quality. The model assumes diminishing returns; beyond certain thresholds, increasing resources, effort, or AI yields smaller gains. AI tools improve efficiency and accuracy, while also reducing misconduct, thus helping to optimise research outcomes. To quantify these relationships, we introduce elasticities εR and εE, representing the proportional contributions of resources, effort, and AI to research quality, respectively. Therefore, we model research quality as a function of resources, effort, and the technological factor as follows:

$$Q = A^{\gamma} R^{\alpha} E^{\beta}, \tag{1}$$

such that $\gamma, \alpha, \beta \in [0,1]$, satisfying the condition that $\gamma + \alpha + \beta = 1$ always. We also assume that the initial values $A(0), R(0), E(0) \geq 0$. Through this model, we capture the interactions between these variables and can assess how changes in resources, effort, and technological enhancement influence the overall research quality. The condition that $\gamma + \alpha + \beta = 1$ enforces constant returns to scale, meaning that a proportional increase in all inputs results in a proportional increase in research quality. The model assumes that increased resources, effort, and AI use improve research output without excessive inflation. This helps evaluate how different inputs affect research quality and integrity, guiding supervision policies.

To begin, we assess the impact of AI on research quality by comparing expected and actual research output within our proposed model. By analysing cases where AI leads to increased retraction rates or diminished originality, we can determine appropriate intervention strategies, such as adjusting AI usage policies or modifying academic integrity thresholds. Research quality follows the functional form given in (1), and if AI does not influence research quality, we define the expected research quality in AI's absence ($\gamma = 0$) as:

$$Q^{*} = R^{\alpha} E^{\beta}, \tag{2}$$

where $Q^{*} \leq Q$ represents the baseline research quality determined solely by resources and effort, and $\alpha + \beta = 1$. The difference between actual and expected research quality allows us to isolate the impact of AI:

$$\Delta Q = Q - Q^* = A^\gamma R^\alpha E^\beta - R^\alpha E^\beta, \tag{3}$$

Rearranging, we obtain:

$$\Delta Q = R^\alpha E^\beta (A^\gamma - 1). \tag{4}$$

The model is not mathematically tractable if $A = 1$, hence we consider the case when $A > 1$, then $\Delta Q > 0$, suggesting that AI enhances research output, potentially by improving access to information, automating writing assistance, or refining data analysis. Conversely, if $0 < A < 1$, then $\Delta Q < 0$, implying that AI negatively impacts research quality, possibly due to overreliance on automated systems, increased plagiarism, or decreased originality in academic work. In Figure 1, we demonstrate the nature of the trajectory of $\Delta Q$ when $A > 1$ revealing a monotonically increasing graph confirming that higher AI use tends to correlate with greater deviations in research quality relative to baseline expectations. This framework allows for the empirical measurement of AI's effect on research quality by estimating the deviation of actual output from expected output.

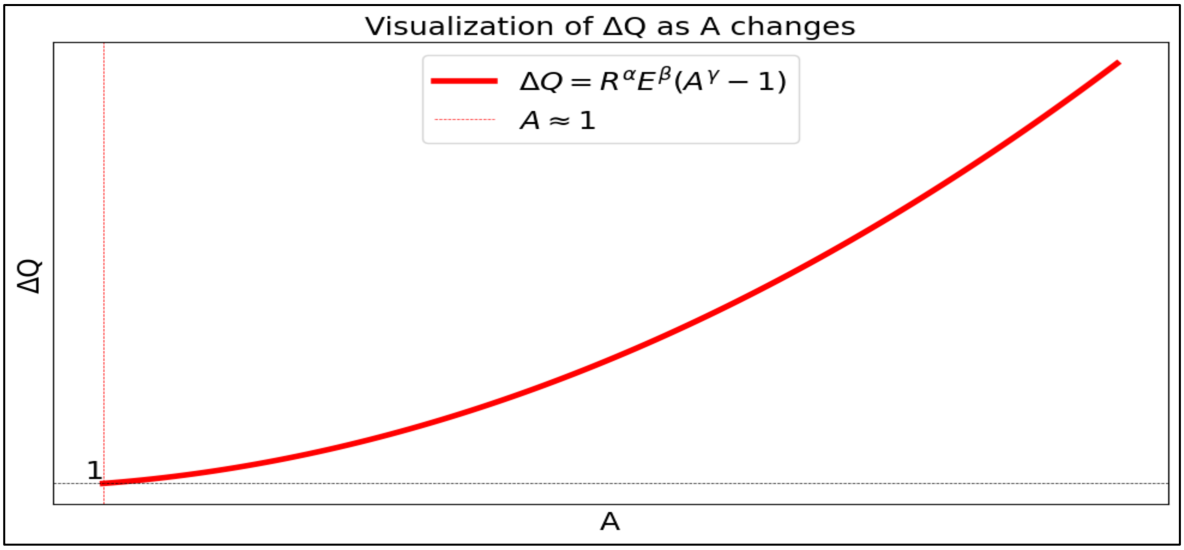



***Figure 1:*** *Graph of $\Delta Q$ as $A$ changes for a constant $\gamma > 0$: The symbolic graph indicates a positive correlation between $A$ and $\Delta Q$, suggesting that higher values of $A$ lead to a greater increase in $\Delta Q$.*

We derive the first-order conditions (FOCs) to determine the optimal allocation of AI, resources and effort for maximizing research quality. Given the research output function in Equation (1), we take the partial derivatives with respect to $A, R$ and $E$ to analyse how research quality responds to changes in these inputs. The partial derivative with respect to $A$ is

$$\frac{\partial Q}{\partial A} = \gamma A^{\gamma-1} R^\alpha E^\beta. \tag{5}$$

Setting this equal to zero would imply an optimal AI level, but since AI usage is typically a decision rather than a resource that can be infinitely adjusted, we must incorporate its cost to make this equation meaningful. We assume that there is a cost associated with AI usage, denoted as $c(A)$ a cost per unit of AI. For simplicity, we assume a linear function $c(A) = kA$. Resources, which include expenses related to AI subscriptions, books, databases, computational tools, and AI-driven research assistants, have a cost $r$. Investing in more resources incurs higher costs, which must be balanced against their contribution to research quality. We also assume that the cost per unit of effort is $w$, representing the opportunity cost of time and energy expended by students and supervisors in conducting research. This can be interpreted as workload, time spent on revisions, or even financial costs such as research grants. The optimisation problem now becomes:

$$\max_{A,R,E}[Q - (rR + wE + c(A))]. \tag{6}$$

Taking the first-order condition of Equation (6) with respect to $A$ from (1) and (5), we get that

$$c'(A) = k = \gamma A^{\gamma-1} R^{\alpha} E^{\beta}. \tag{7}$$

This equation states that the marginal benefit of AI on research quality should equal its marginal cost. If AI is costly or leads to unintended negative consequences (such as increased retractions or loss of originality), adjusting its usage becomes necessary. If the marginal benefit of AI exceeds $k$, (that is, $\gamma A^{\gamma-1} R^{\alpha} E^{\beta} > k$ ), increasing $A$ raises net quality; if below $k$, decreasing $A$ is optimal. If the optimal AI level is exceeded (that is, $A$ is too high), it may indicate that AI is replacing genuine research effort rather than complementing it, necessitating intervention.

For resources, the first-order condition is:

$$\frac{\partial Q}{\partial R} = \alpha A^{\gamma} R^{\alpha-1} E^{\beta} . \tag{8}$$

This result implies that when $\alpha$ is high, resources such as library access, technological tools, and academic databases play a pivotal role in improving research quality. If $R$ It is too low; investing in these resources becomes essential to maintain high research standards. Conversely, when $\alpha$ is low, the impact of additional resources is minimal, indicating that increasing $R$ does not substantially enhance research quality. Instead, other factors such as researcher effort or the AI play a more dominant role.

For effort, the first-order condition is:

$$\frac{\partial Q}{\partial E} = \beta A^{\gamma} R^{\alpha} E^{\beta-1}. \tag{9}$$

This result suggests that if $\beta$ is high, research quality is more dependent on human effort. A low level of effort, particularly due to an over-reliance on AI tools, may compromise academic rigour and originality. To determine the optimal balance between resources, effort, and AI usage, we introduce a cost function that includes the cost of resources ($rR$) and the cost of effort ($wE$). The objective is to maximise net research quality:

$$\max_{A,R,E}[Q - (rR + wE + c(A))]. \tag{10}$$

Taking the first-order conditions for this optimisation problem, we obtain that:

$r = \alpha A^{\gamma} R^{\alpha-1} E^{\beta}$ and $w = \beta A^{\gamma} R^{\alpha} E^{\beta-1}$.

These conditions establish equilibrium points where the marginal contribution of resources and effort to research quality equals their respective costs.

We introduce a utility framework that incorporates a temporal adjustment mechanism, recognising that the perceived value of research output evolves over time. As AI becomes more integrated into research practices, its marginal utility diminishes due to factors such as over-reliance, reduced originality, and ethical concerns. We introduce a discounting factor, $\delta \in [0,1]$, which adjusts the utility derived from each additional AI-enhanced research output, where $n(t)$ is the number of published articles on AI. The student's utility at the time $t$ follows:

$$U(t) = U(Q_{\infty}) + \delta^{n(t)}[U(Q_0) - U(Q_{\infty})] \tag{11}$$

where $U(Q_0)$ and $U(Q_{\infty})$ represent the utility levels when AI use is minimal and at its maximum saturation, respectively and $Q_{\infty} = \max_{A,R,E}[Q - (rR + wE + c(A))]$ as given in Equation (10). This formulation captures the transition from initial AI adoption, where its benefits are pronounced, to a scenario where excessive reliance leads to diminishing returns. The discounting effect ensures that AI usage is dynamically regulated, preventing overuse that might compromise research integrity. By fitting this model to AI publication data, we obtained the following fitted parameters $\delta = 0.8123$, $U(Q_0) = 7.26 \times 10^{-16}$, $U(Q_{\infty}) = 1$ and an $R^2$ value of $0.8735$ indicating a solid fit to the observed trends.

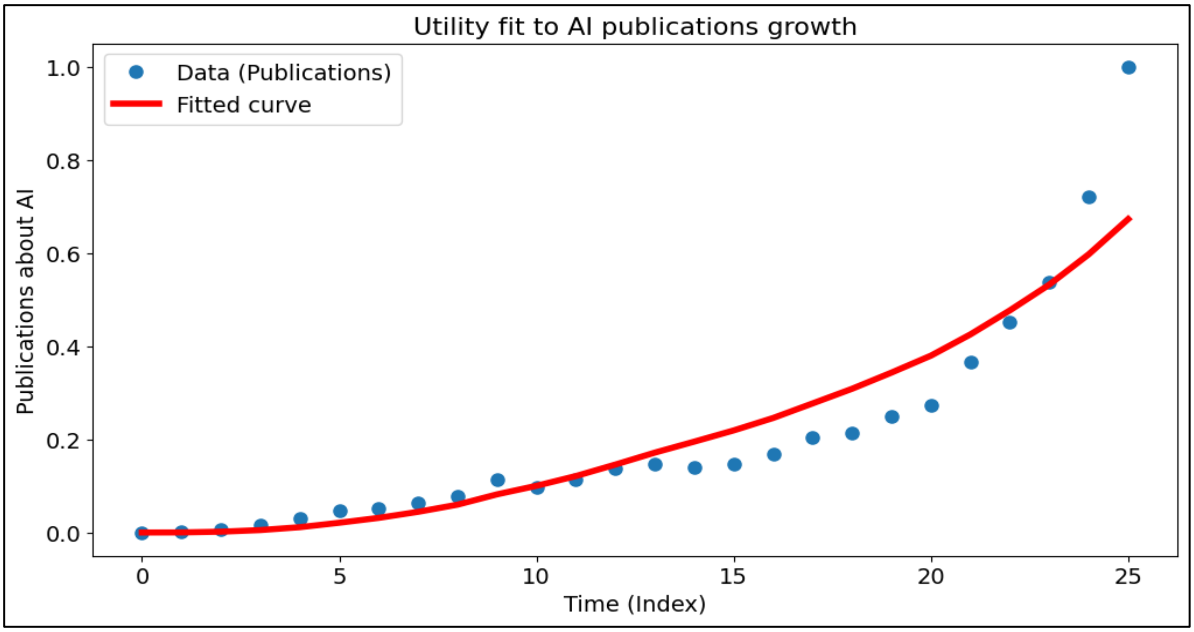



***Figure 2:*** *Fitted utility curve modelling the evolving impact of AI-enhanced research over time, with estimated initial values for* $\delta, U(Q_0), U(Q_{\infty})$ *being 1, 0, and 0.95, respectively*

The AI publications data is a proxy for AI awareness growth and spread annually. The results in Figure 2 suggest that while the initial perceived utility of AI-enhanced research was almost negligible, it increased rapidly with adoption and then gradually levelled off as the number of AI-related publications grew. The long-term utility stabilises at a positive value of 1, which means that even in a saturated research environment, AI maintains a lasting, meaningful contribution. However, its incremental benefits diminish over time, emphasising the importance of thoughtful and measured AI adoption in research to preserve both the value of AI and the integrity of the research field as it continues to evolve. For the rest of the study, without loss of generality, we work with the maximum utility 1, which is negligible as a factor and does not affect the model's outcomes.

### 2.2 The heterogeneous case

To model the heterogeneity of AI use within postgraduate research with maximum utility, we introduce an AI adoption distribution function that accounts for variations across researchers, disciplines, and institutions. This approach deviates from the assumption of uniform AI utilisation by acknowledging that AI integration into research workflows is influenced by factors such as familiarity with AI tools, institutional policies, ethical considerations, and field-specific norms. To understand individual student behaviour, we introduce distinct researcher types indexed by $i = 0,..,n$, where each researcher $i$ uses AI at a distinct level $A_i$. The aggregate research quality across all researchers is then expressed as:

$$Q_{agg} = \sum_i A_i^\gamma R^\alpha E^\beta \ . \tag{12}$$

In cases where AI heterogeneity arises due to institutional access, career incentives, or disciplinary norms, we define AI use for researchers $i$ as $A_i = A_0 + \theta_i A_0$ where $A_0$ represents baseline AI access, and $\theta_i$ captures individual deviations due to external factors or preferences. If excessive AI use compromises research integrity, institutions may restrict its application in specific areas. Conversely, if underuse results from limited access or insufficient training, policies may prioritise AI literacy and equitable access to resources.

Building on the understanding that AI adoption varies across different research tasks, we model the effect of AI use on research quality by treating research output as a collection of individual research articles within a broader academic research space, denoted as $B$. The articles within the research space $B$ are put into classes $b_1, b_2, \dots, b_j$, where each $b_j$ corresponds to a specific aspect of research quality, such as novelty, proper citation practices, depth of analysis, and evaluative reasoning. For simplicity, we assume that each $b_j$ is independent of the rest, hence, coupling or cross terms are not considered. To account for AI's influence, we assume a baseline set of AI usage levels $A_0 = \{z_1, z_2, \dots, z_k\}$ where each $z_i$ represents the responsible, acceptable level of AI use that supports quality research in the corresponding component $b_i$ without undermining academic standards. However, the actual AI use intensities, denoted as $\theta_i A_0 = \{s_1, s_2, \dots, s_k\}$ can deviate from these baselines, potentially introducing ethical risks to research quality. To quantify the deviation from responsible AI use, we define the difference between actual AI use and the baseline levels in each $b_j$ as:

$$\{t_{11}, t_{12}, \dots, t_{kj}\} = \{s_{11}, s_{12}, \dots, s_{kj}\} - \{z_{11}, z_{12}, \dots, z_{kj}\}. \tag{13}$$

Each research component $b_j$ exhibits two functions: $\Gamma_j(t_{kj})$, which captures the positive contribution of AI within an acceptable range of use, and $d_j(t_{kj})$, which quantifies the potential negative effects of AI overuse. These negative effects may include issues such as increased plagiarism, reliance on AI-generated content, reduced originality, or weakened engagement with research materials. To prevent AI use from exceeding acceptable thresholds, we introduce a damage limit $\varphi_j$ for each research component $b_j$, enforcing the condition:

$$\sum_{t_{kj} \in T} d_j(t_{kj}) < \varphi_j \quad \text{for all } k. \tag{14}$$

where $T$ represents the set of AI use deviations beyond the baseline. If the cumulative negative impact $D_j(t)$ for any component exceeds $\varphi_j$, research quality is considered compromised within that dimension.

This reasoning is illustrated in the model flow in Figure 3 that follows:

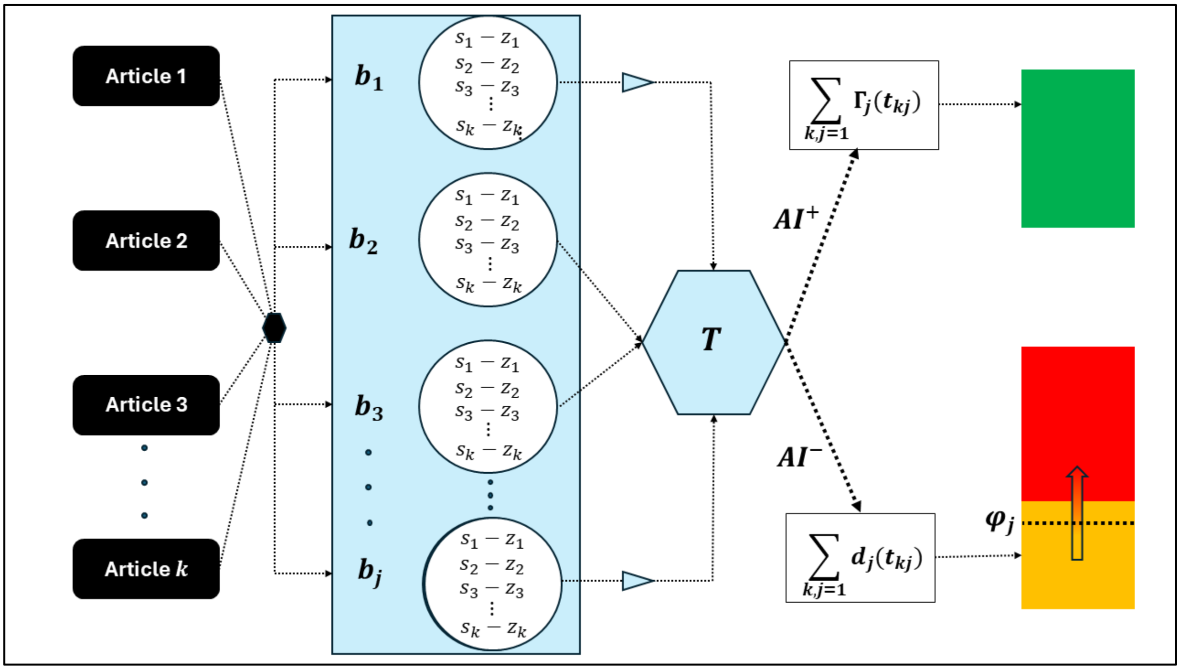



***Figure 3:*** *An illustration of how the model classifies and quantifies positive and negative gains by considering the research components $b_j$. The arrows for $AI^+$ and $AI^-$ denote positive and negative benefits of AI use, respectively*

By combining these ideas, we can express the total research quality as a sum of overall quality dimensions. Specifically, the overall research quality can be written as

$$Q_{agg} = \sum_k^K \left(\sum_{i=1}^N \left(\sum_i A_i^\gamma R^\alpha E^\beta \left[\Gamma_j(t_{kj}) - d_j(t_{kj})\right]\right)\right). \tag{15}$$

The inner summation over $i$ aggregates the contributions from all researchers based on their individual AI use $A_i$ (modulated $by\ A_i^{\gamma}$), research rigour $R^{\alpha}$, and student effort $E^{\beta}$. The function $\Gamma_j(t_{kj})$ represents the beneficial contribution of AI to the $k$-th component when operating at its baseline level $z_i$, while $d_j(t_{kj})$ captures the detrimental impact arising from deviations $t_{kj}$ beyond the baseline. The cumulative damage for each research component $b_j$ is calculated based on the deviations $t_{kj}$ and if the total negative impact $d_j(t_{kj})$ for any dimension exceeds its threshold $\varphi_j$, we consider that component to have experienced a compromise in research quality, allowing for the management of AI adoption in research.

The plagiarism tolerance $P$ is directly related to the acceptable threshold for negative effects arising from AI use. In this context, we define damage tolerance $\varphi_j$ as the upper limit for damage within each research component $b_j$, and plagiarism tolerance $P$ as the threshold that ensures research quality remains uncompromised. The relationship between damage tolerance and plagiarism tolerance can be expressed as follows:

$$d_j(t_{kj}) \leq \varphi_j \leq P \tag{16}$$

This means that the cumulative damage resulting from AI overuse should not exceed the plagiarism tolerance $P$, which serves as the limit that ensures the integrity of research is maintained. When introducing plagiarism tolerance $P$ into the overall research quality $Q_{agg}$, we must account for the effect of AI overuse on the research quality. The cumulative damage for each research component $b_j$ influences the research quality, as seen in Equation (15). The role of plagiarism tolerance $P$ is to prevent the cumulative damage from exceeding the acceptable limits. This ensures that $d_j(t_{kj})$ does not exceed the threshold at which research quality would be compromised. If $d_j(t_{kj})$ exceeds $P$, strict intervention is required to mitigate the overuse of AI and prevent further damage to research quality(for example, the faculty can reject a thesis submission). If $d_j(t_{kj}) \leq \varphi_j \leq P$, the research quality remains unaffected by AI overuse, and the total quality $Q_{agg}$ is not compromised.

To account for plagiarism tolerance in adjusting the overall research quality $Q_{agg}$, a correction factor is introduced into the formula for total quality. This reflects the impact of the acceptable damage due to AI overuse:

$$Q_{agg} = \textstyle\sum_k^K \left(\sum_{i=1}^N \left(\sum_i A_i^{\gamma} R^{\alpha} E^{\beta} \left[\Gamma_j(t_{kj}) - \min(d_j(t_{kj}), P)\right]\right)\right). \tag{17}$$

The minimum function ensures that if the damage $d_j(t_{kj})$ exceeds $\varphi_j$, the damage is capped at $P$ to prevent further degradation of research quality. If the damage is less than or equal to $\varphi_j$, it remains as is, contributing to the overall quality without requiring intervention. The inclusion of plagiarism tolerance $P$ in the formula for $Q_{agg}$ ensures that AI overuse, particularly regarding plagiarism or reduced originality, does not undermine research quality beyond an acceptable threshold. This threshold is determined by $P$Intervention occurs when damage exceeds this value, adjusting the total quality to account for the negative impacts of AI. Therefore, $P$ sets the upper limit for the damage $d_j(t_{kj})$ that is acceptable to maintain research quality, and the total research quality $Q_{agg}$ is dynamically adjusted to prevent AI overuse from undermining important aspects such as originality, rigor, and academic integrity.

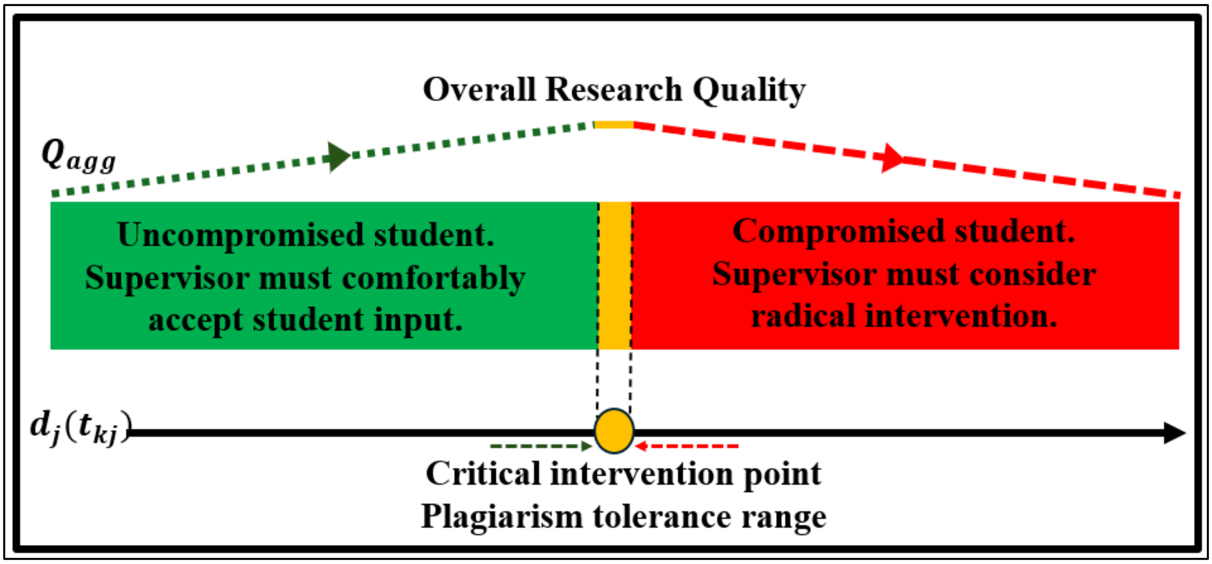



***Figure 4:*** *A schematic illustration of the effect of AI use and abuse. Initially students are presumed to be responsible and at the end, the damage calls for radical intervention*

Figure 4 presents two scenarios: in the case of an uncompromised student ($d_j(t_{kj}) < \varphi_j$), the supervisor can comfortably accept the student's input, which contributes positively to the aggregate research quality ($Q_{agg}$). Conversely, if the student is compromised ($d_j(t_{kj}) > P$), the supervisor may need to consider more drastic measures, as indicated by the red zone, which poses a threat to research quality. The intervention point serves as a critical boundary (when $\varphi_j < d_j(t_{kj}) < P$), prompting the supervisor to decide whether intervention is necessary based on the student's performance within the plagiarism tolerance range. This visual representation reinforces the delicate balance between student autonomy and the need for oversight in ensuring high-quality research output.

### 2.3 Modelling with data

This study relies on two core datasets, as shown in Figure 5: the number of peer-reviewed published articles and the number of retracted articles. Each dataset is selected based on its alignment with the central aim of analysing the relationship between research capacity, research output, and research integrity. The temporal scope of all datasets spans from 2000 to 2024, providing a 24-year window that enables the capture of long-term structural and systemic trends in global scientific production.

The publication and retraction datasets are derived from the Scopus database. Scopus is selected due to its broad disciplinary coverage, standardised indexing practices, and global reach, making it a suitable source for analysing trends in scholarly output across time and regions. Scopus includes metadata on authorship, affiliations, and document types, which supports disaggregation and further analysis. Its reliability and comprehensiveness make it an appropriate proxy for global research activity and integrity.

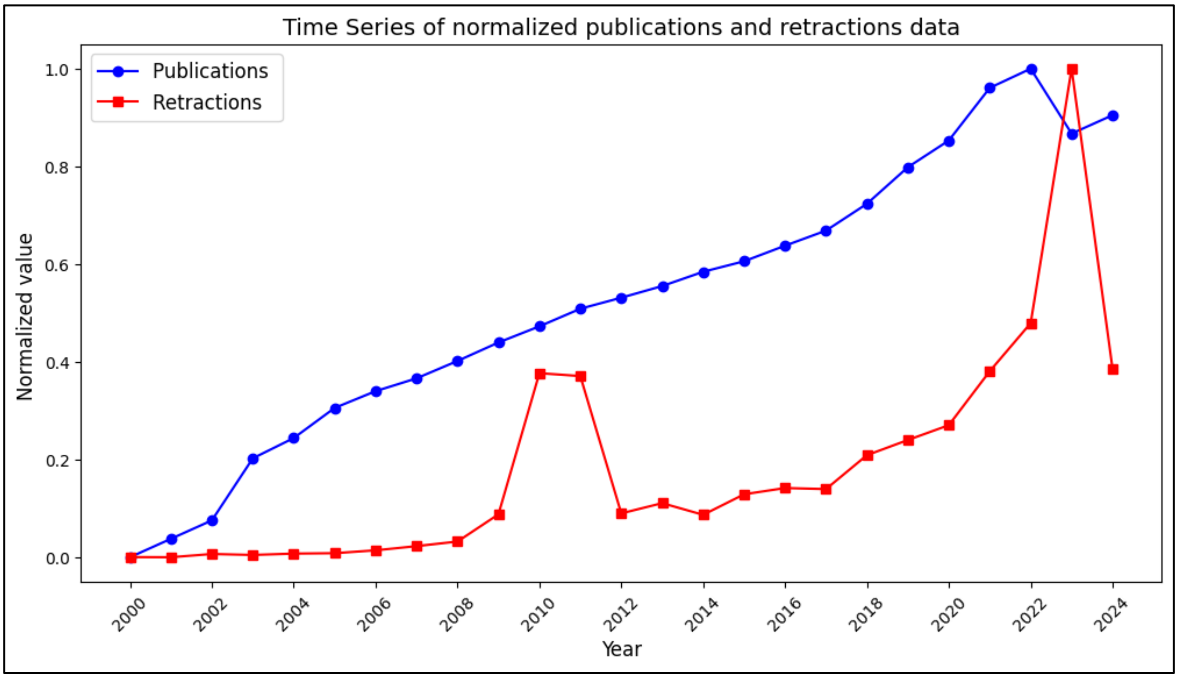



***Figure 5:*** *The plot illustrates the evolving relationship among the three variables. A notable rise in retractions is observed after 2011–2012, following earlier increases in publications, suggesting a possible quality–quantity trade-off*

Retractions are formal notices that a publication has been withdrawn due to error, misconduct, or other violations of publication ethics. While retractions do not capture all dimensions of research integrity, they are one of the few standardised and publicly traceable signals of compromised scientific quality. Including this metric allows the study to empirically assess how the growth of research output may affect the prevalence of integrity failures. The global scope of retraction data is intentional, as research misconduct and correction practices are not limited to a specific country or system, and cross-national trends in retractions are evidence of systemic pressures affecting scientific rigour.

We normalise all datasets using min-max scaling ([0,1]) to control for systemic factors like population growth and expansion in higher education, ensuring trends reflect real shifts in research behaviour, not structural changes. This allows proportional comparisons across time, revealing whether retraction trends persist when adjusted for rising publication volumes. Normalisation is essential to accurately assess the link between research output and integrity. The datasets feed into a utility function that models how increased activity may trade off with quality, incorporating time lags to capture delayed effects. This approach ensures methodological rigour and interpretive clarity.

## 3. Presentation of Results

We provide a comprehensive analysis of research integrity risks, including the estimation of retraction-based risks and the identification of key thresholds. Lag analysis highlights critical intervention points, while the quantification of plagiarism tolerance underscores the impact of AI misuse on academic quality. Additionally, we explore the dynamics of institutional enforcement and model the time-to-failure of research quality under sustained AI-related misconduct.

### 3.1 Estimating retraction-based integrity risks

To indirectly estimate the impact of AI misuse on research integrity, we define an AI misuse impact index, denoted as $AI_x(t)$. This index is based on the assumption that retractions are primarily due to some unethical conduct on the part of the authors, and it is calculated as the ratio of the normalized number of retractions to the normalized number of publications in a given year, with a small positive constant $\epsilon$ (set to 0.001) added to the denominator to avoid division by zero in years with negligible publication output. The formula is expressed as:

$$AI_x(t) = \frac{\text{Number of retractions}}{\text{Number of Publications} + \epsilon} \tag{18}$$

A rising value of $AI_x(t)$ indicates an increasing integrity risk relative to research output, potentially reflecting the growing overuse or misuse of AI tools in academic research. Analysing the data, we observe that the $AI_x(t)$index remains negligible in the early 2000s, with values such as 0.0878 in 2002 and 0.0230 in 2003. However, a sharp increase appears around 2010, where the index jumps to 0.7973, signalling a substantial rise in integrity risks. This trend continues, with the index reaching 0.3963 in 2021 and 0.4792 in 2022. Notably, it peaks dramatically at 0.7975 in 2023, suggesting a critical point where AI misuse may have exceeded acceptable thresholds, likely corresponding to the plagiarism tolerance $P$ in the model.

This pattern in Figure 6 aligns closely with the framework's assumptions that before 2010, AI's influence on academic norms appears minimal, with low integrity risks. Between 2010 and 2020, rising AI adoption coincided with steadily increasing retraction rates relative to publications, indicating growing concerns. From 2020 to 2023, the proxy index spikes rapidly, suggesting significant overuse or unethical reliance on AI tools in the model where integrity risks surpass tolerable limits. The finding that $AI_x(t)$ exceeds 1.0 in 2023 strongly suggests that the cumulative damage from AI misuse likely breached both the damage limit $\phi_j$ and plagiarism tolerance P, necessitating strict intervention to preserve research integrity. This pattern aligns closely with the framework's assumptions that before 2010, AI's influence on academic norms appears minimal, with low integrity risks. Between 2010 and 2020, rising AI adoption coincides with steadily increasing retraction rates relative to publications, indicating growing concerns. From 2020 to 2023, the proxy index spikes rapidly, suggesting significant overuse or unethical reliance on AI tools in the model where integrity risks surpass tolerable limits.

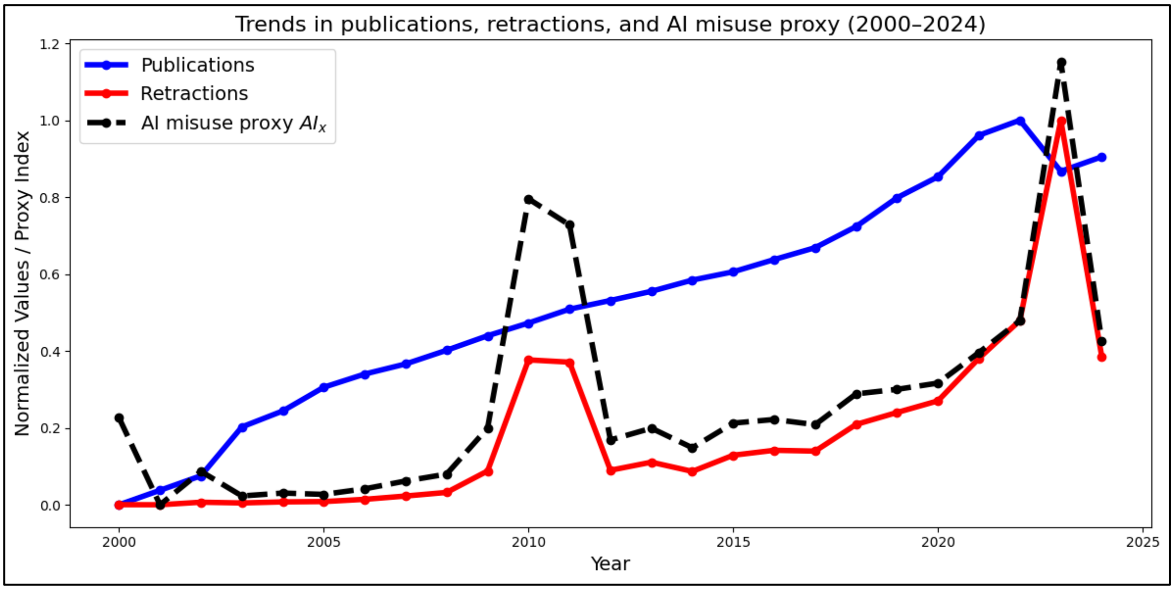



***Figure 6:*** *Normalized trends in publications, retractions, and AI misuse proxy (2000–2024). Notice the sharp AI proxy spikes post-2010 and peaking in 2023, indicating rising integrity risks alongside AI adoption*

### 3.2 Lag analysis of retraction rates and thresholds estimation

Using time series correlation and lag analysis, we explore whether spikes in retraction rates follow increases in publication output. The results show that retraction rates are most strongly correlated with submissions one year later (Pearson correlation coefficient of 0.8576 at lag 1), indicating that unethical AI use, such as plagiarism or manipulation, manifests within a year. As the lag increases, the correlation weakens (0.6207 at lag 2 and 0.4501 at lag 3), supporting the hypothesis that AI-driven misconduct has a relatively short-term impact on retraction decisions.

Given this observation, which is illustrated in Figure 7, it is essential to define thresholds for intervention based on the most relevant time window and, more specifically, the 1-year lag period in which correlations are strongest. We use piecewise linear models rather than nonlinear alternatives like exponential or quadratic functions to identify explicit threshold points where the relationship between normalised publications and retractions changes behaviour. Piecewise linear models estimate these breakpoints directly, making them well-suited for detecting shifts associated with integrity risk thresholds in our theoretical framework. Nonlinear models can capture general trends but do not provide interpretable points of change. The data exhibits distinct phases of growth that align naturally with segmented linear behaviour, supporting the use of this model form.

We apply two models: a 2-segment and a 3-segment piecewise linear fit. The 2-segment model, with an R-squared value of 0.5446, identifies breakpoints at 0.0 and 0.6585 publications, capturing general retraction acceleration. The 3-segment model, with a higher R-squared value of 0.7517, identifies breakpoints at 0.0, 0.8525, 0.8615, and 1.0, offering a more detailed analysis. We use the 3-segment model, as it provides a more accurate reflection of retraction rate shifts, identifying the $P$-threshold where retraction rates significantly increase, and interventions may be needed to protect research integrity. In Figure 7, the spline reveals whether changes in retraction rates occur abruptly at the breakpoints or follow a continuous, gradual trend, helping assess if the segmented thresholds reflect real inflexion points or oversimplify a nonlinear relationship. These inflections represent key moments when the system's response to publication volume changes, suggesting areas where the retraction rate might either increase or slow down based on trends in AI misuse.

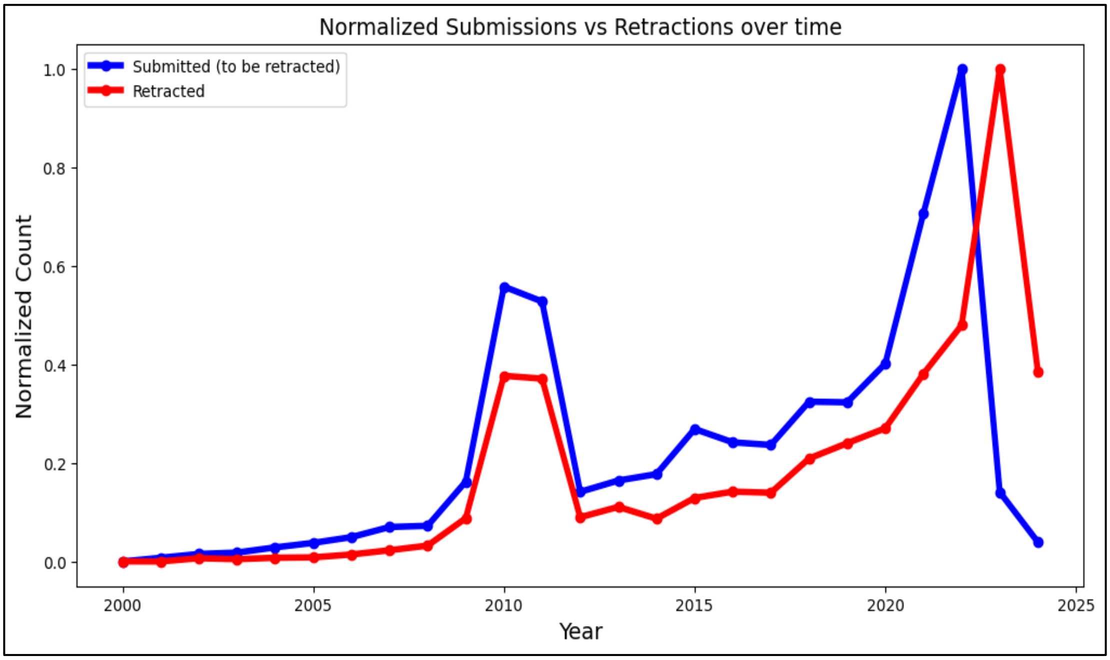



***Figure 7:*** *A comparison of the trends in the number of submissions and retractions from 2000 to 2024*

We conduct a piecewise linear analysis to identify thresholds in the relationship between normalised publications and retractions, aiming to pinpoint when retraction rates accelerate.

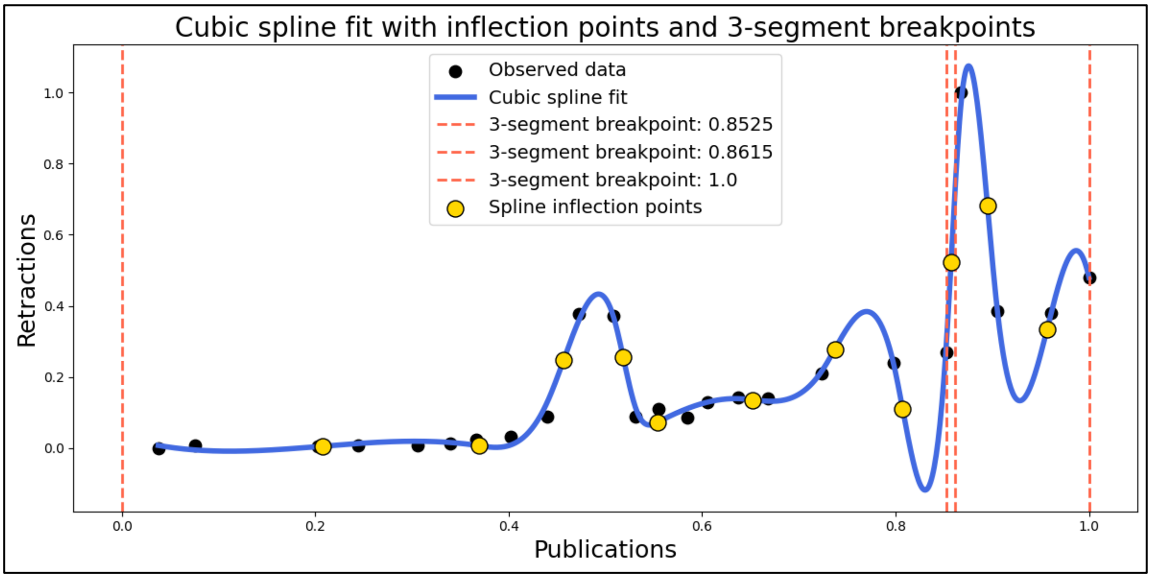



***Figure 8:*** *Comparison of a cubic spline fit to normalised publication and retraction data, overlaid with vertical lines representing the 3-segment model breakpoints*

### 3.3 AI-induced ethical risks and plagiarism tolerance thresholds

Our model accounts for varying AI adoption across disciplines, institutions, and researchers by introducing individual AI use levels (Equation 12). Using longitudinal data, we calculate a marginal damage ratio, which links retractions to increased scholarly output. We also determine the plagiarism tolerance value by calculating an annual AI misuse impact index, reflecting misconduct relative to academic productivity.

We determine the plagiarism tolerance threshold $P$ by examining the AI misuse impact index, $AI_x(t)$and identifying the first year in which it exceeds a selected inflection point value derived from the data. This inflection point represents a moment where the relationship between retractions and publication volume begins to shift noticeably, signalling growing systemic risk. The corresponding $AI_x(t)$value in that year is then set as P. If no year surpasses the selected inflection point, the highest observed $AI_x(t)$value is used instead. This empirically derived tolerance acts as a cap on AI misuse's damage in subsequent simulations, limiting its degrading effect on overall research quality. Rather than being a theoretical parameter, $P$ is a data-driven indicator of when misconduct begins to accelerate disproportionately relative to academic output, marking a threshold where research integrity is at risk.

To reflect this in the heterogeneous model, we extend the concept of quality to include specific components of research such as originality, citation practices, and rigorous analysis (each represented by subspaces $b_j$). Each of these components has a baseline AI use $z_k$, a beneficial range $\Gamma_j(t_{kj})$ and a damage function $d_j(t_{kj})$, as given in Equations (13). The total quality is adjusted by the minimum of the damage and the threshold $P$, as shown in Equation (15), a formula that operationalises the plagiarism tolerance $P$ as a regulatory boundary. If AI overuse in any dimension $b_j$ causes damage $d_j(t_{kj})$ to exceed $P$, the system triggers intervention to cap further degradation. When damage is within tolerance (see Equation (16)), no correction is needed. This creates a natural "stop" mechanism on the decline of academic quality, ensuring that research output remains credible even in an AI-augmented environment.

To do this, we calculate the marginal damage from academic retractions relative to the growth in  and articles to estimate a plagiarism tolerance threshold $P$. We then simulate overall academic quality $Q_{agg}$ under two scenarios: one where damage is unchecked, and another where it is capped at $P$, to illustrate how enforcing a threshold preserves systemic research integrity. We assign equal weights $\gamma = \alpha = \beta = \frac{1}{3}$ to reflect the assumption that , publications, and effort equally add to academic quality in the absence of detailed empirical data on their relative impacts. This approach ensures a balanced contribution while maintaining neutrality in the absence of bias toward any one factor. We set $E = 1$ to normalise effort, assuming it remains constant across years, which is reasonable since academic supervisors with tenure do not have a high turnover rate.

This simplification isolates the effects of  publications, leaving focus specifically on how damage (via retractions) impacts quality, with and without the plagiarism tolerance threshold. To get Figure 8, we first compute year-over-year changes in retractions, and published articles, then divide the retraction change by the article change to get marginal damage ratio. Using these ratios, we identify the threshold $P$ where damage exceeds a critical value and simulate academic quality $Q_{agg}$ both with and without capping damage at $P$, showing the impact on system stability through a comparative plot. In the unconstrained scenario, academic quality $Q_{agg}$ erodes rapidly as retractions accelerate past sustainable levels. However, when we enforce the plagiarism tolerance $P$, quality is preserved, and the system stabilises, demonstrating that this threshold has not only mathematical elegance but also real-world utility. A negative aggregate research quality score indicates that AI is being overused to the point where its associated damage (such as plagiarism or loss of originality) outweighs its benefits. This suggests that either the plagiarism tolerance $P$ is too low or the damage function $d_j(t_{kj})$ escalates too rapidly, signalling systemic quality degradation and the need for immediate intervention.

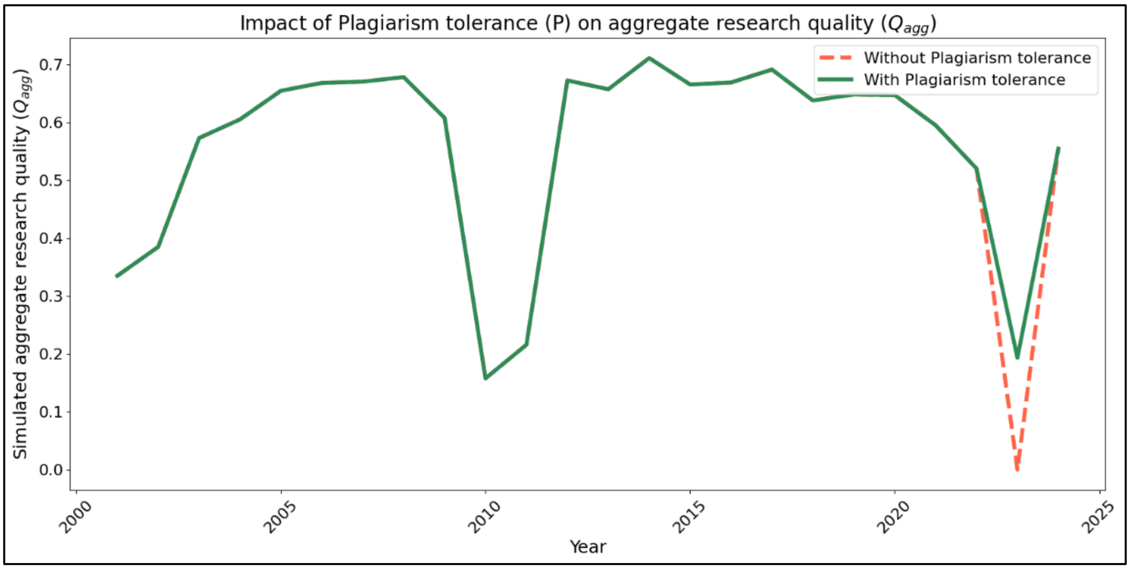



***Figure 8:*** *Simulated academic quality ($Q_{agg}$) with and without plagiarism tolerance*

The gap between the two curves (green and red) in Figure 8 reflects the potential benefit of enforcing a tolerance threshold in controlling the erosion of research integrity

### 3.4 Quantifying institutional plagiarism enforcement

To further contextualise $P$, we map it to familiar academic integrity tools, such as Turnitin which is simply the acceptable Similarity index at an institution $S$. We begin by considering the institutional plagiarism tolerance as a function of both the institution's stated standards and its actual performance in curbing misconduct. The ratio $\frac{AI_x(t)}{S}$ compares the actual impact of AI misuse $AI_x(t)$, which reflects the institution's acceptable plagiarism score, to the declared tolerance threshold ($S$), which is measured by the ratio of retractions to publications. This ratio helps assess how well an institution's policies align with enforcement. A ratio greater than 1 suggests effective oversight or low misconduct. A ratio near 1 indicates misconduct approaching the declared threshold, requiring stronger oversight. A ratio less than 1 shows the retraction rate exceeds the declared tolerance, signalling inadequate enforcement. We introduced earlier the multiplier $P$, which represents the system-wide empirically observed plagiarism tolerance threshold. This threshold is based on the broader trends in AI misuse, where research quality begins to degrade when misconduct rises above a critical level. By multiplying the ratio $\frac{AI_x(t)}{S}$ by $P$, we scale the institution's tolerance dynamically, aligning it with the system-wide risk of misconduct. This adjustment ensures that institutional thresholds are sensitive to both the institution's own performance and the broader academic environment. Thus, we arrive at the final formula for the adjusted institutional plagiarism tolerance (the ideal one, $S_{ideal}$):

$$S_{ideal} = P \times \frac{AI_x(t)}{S} \tag{19}$$

This formula provides a dynamic and data-driven approach to adjusting an institution's plagiarism tolerance based on both its own conduct and broader systemic trends in academic misconduct.

The ratio between the actual retraction rate and the institution's acceptable plagiarism tolerance $\frac{AI_x(t)}{S}$, serves as a diagnostic indicator of policy enforcement and research integrity management. This metric captures the gap between declared academic standards and real-world accountability. For example, if an institution claims an acceptable plagiarism threshold of 20% while the actual retraction rate is 0.1% (equivalent to 1 retraction per 1,000 publications), the resulting ratio would be 0.005. Such a value suggests that the institution's retraction rate is only 0.5% of its stated institution's plagiarism score ($S = 20\%$), suggesting either weak enforcement, low actual misconduct, or underreporting. When the ratio exceeds 1, it signals that misconduct is rising beyond what the institution claims to tolerate, indicating a critical gap between policy and practice that demands immediate attention.

### 3.5 A probabilistic integrity decay framework

Our model extends to incorporate temporal dynamics that capture how research quality degrades progressively due to sustained AI misuse. For each research quality dimension $b_1, b_2, \dots, b_j$, deviations from the ideal AI baseline $z_{kj}$(quantified as $t_{kj} = s_{kj} - z_{kj}$) introduce incremental damage. These deviations, as defined in Equation (19), manifest as erosion in originality, increased plagiarism, or weakened analytical depth. The cumulative damage to the dimension $b_j$ over $t$ years,

$$D_k(t) = \sum_{t_j \in T} d_j(t_{kj}) < \varphi_j, \qquad (20)$$

triggers a need for intervention in that dimension if it surpasses the threshold $\varphi_k$. At the system level, total accumulated damage

$$C_t = \sum_{k=1}^{K} D_k(t) \qquad (21)$$

is measured against a plagiarism tolerance threshold $P$. System failure occurs when $C_t > P$, with the time to intervention marking the earliest point at which corrective actions (consider as examples policy reforms or audits) become necessary. This framework dynamically links AI misuse to quality decay, enabling real-time monitoring of integrity breaches. For damage accumulation, we assume each article's contribution

$$D_i - D_{i-1} = S_i \psi(D_{i-1}), \qquad (23)$$

where $S_i$ is the $i$-th article's vulnerability and $\psi(D_{i-1})$ models ethics decay from prior damage. With $\psi(D) = \omega D$ (assume linear decay for simplicity), solving $\sum_{i=1}^{k} S_i$ yields:

$$\ln\left(\frac{D_k}{D_0}\right) \sim \mathcal{N}(\mu_A, \sigma_A^2), \qquad (24)$$

revealing log-normally distributed damage. This shows minor, initially imperceptible deviations compound exponentially, culminating into abrupt ethics failure once $C_t > P$. Consequently, the time to intervention becomes a probabilistic metric, emphasising that integrity breaches often emerge nonlinearly, necessitating proactive monitoring long before thresholds are visibly exceeded.

The risk of failure (when $C_t > P$) can now be probabilistically modelled using the cumulative damage distribution:

$$\mathrm{Prob}(C_t > P) = 1 - \Phi\left(\frac{\ln\left(\frac{P}{C_0}\right) - \mu_A}{\sigma_A}\right). \tag{25}$$

This equation allows us to quantify the probability of exceeding the threshold based on the AI adoption behaviour. In this expression, $\Phi(\cdot)$ represents the cumulative distribution function (CDF) of the standard normal distribution. The term $C_0$ stands for the initial level of damage, which we can normalize to 1 for simplicity. The parameters $\mu_A$ and $\sigma_A$ capture the mean and dispersion of AI usage and misuse across researchers, as specified in the AI adoption model. This formulation allows us to probabilistically estimate the risk of breaching the system-wide tolerance threshold based on observed trends in AI misuse and damage accumulation. Hence, we modify Equation (19) to include the probability of exceeding the critical plagiarism tolerance in Equation (25) and get:

$$S_{ideal} = P \times \mathrm{Prob}(C_t > P) \times \frac{AI_x(t)}{S}. \tag{26}$$

This adjustment makes $S_{ideal}$ responsive to the probabilistic risk of exceeding the critical plagiarism threshold, scaling tolerance based on both AI adoption and integrity risk. It ensures dynamic regulation, tightening controls as $\mathrm{Prob}(C_t > P)$ rises.

## 4. Discussion of Findings

Integrating AI in postgraduate supervision boosts research efficiency but introduces costs, supervision time, and academic integrity risks. The optimisation framework (Equation 6) balances AI use, resources, and human effort to maximise quality while minimising costs. It guides decisions on when to use AI or rely on mentorship, ensuring that AI enhances quality without compromising ethics. This framework supports responsible AI integration, safeguarding both academic standards and mentorship.

The plagiarism tolerance threshold $P = 0.797$, with the threshold year being 2009-10, serves as a data-driven intervention point, marking where retractions begin to rise disproportionately relative to scholarly output. This value, empirically derived from the relationship between academic retractions and output, highlights a critical inflection point where misconduct threatens systemic integrity. Rising retractions, especially when reinforced by lagged effects, suggest the need to lower P to prevent broader ethical failures. A predictive model using lagged retraction data can forecast when this threshold be exceeded, signalling the need for corrective action. By integrating this threshold into governance practices, institutions can proactively manage academic risk, maintaining research credibility in the face of evolving AI-driven challenges.

The model's accommodation of heterogeneous AI adoption across disciplines (for the same institution, different students behaviour) represents another important advancement. By incorporating individual-specific AI use levels (as expressed in Equation (12)), the framework recognizes that AI's effects are context-dependent. This allows for more tailored policy interventions rather than relying on uniform, oversimplified assumptions. Complementing this is the dynamic damage capping mechanism, wherein $P$ is enforced as a regulatory boundary as in Equation (16), functioning as a circuit breaker that prevents unchecked degradation of research quality. This feature, illustrated in Figure 3, demonstrates how capping damage at the empirically derived threshold stabilises aggregate research quality ($Q_{agg}$), offering institutions a pragmatic safeguard.

The marginal damage ratio quantifies the cost of AI misuse per unit of academic productivity, linking retractions to publication growth. This framework incorporates probabilistic risk assessment, modelling integrity breaches as log-normally distributed events via Equations (24) and (25). Due to delays in detection and retraction processes, current retraction rates may underestimate misconduct, particularly in fields adopting AI. This emphasises the importance of the plagiarism tolerance threshold $P$ as an early warning indicator of rising systemic risks before retractions accumulate.

In this integrated model, the log-normal distribution emerges as the natural fit for capturing the accumulated damage due to the heterogeneity in AI adoption, where the system's overall quality degradation reflects the skewed distribution of AI reliance. This way, both individual and collective damage from AI use are modelled, with interventions triggered once cumulative damage $C_t$ surpasses a threshold $P$, marking system failure. This log-normal behaviour of cumulative damage implies that while most research systems may remain below the plagiarism tolerance threshold $P$ for an extended period, a smaller subset (due to compounding misuse and prior degradation) can experience a rapid, disproportionate escalation in damage. As a result, the system-wide risk of crossing the threshold $P$ is not gradual but can manifest suddenly and unexpectedly. This underscores the importance of early monitoring and intervention, since once the cumulative damage $C_t$ enters the long tail of the log-normal distribution, the probability of surpassing $P$ increases non-linearly, making recovery more difficult and requiring more severe corrective action.

Despite these strengths, the model also presents critical limitations and challenges. One issue lies in the assumption of constant academic effort ($E = 1$), which normalises researcher input and ignores changing trends such as productivity inflation, burnout, or shifts in time allocation driven by AI tools. A more realistic approach would model effort as a function of AI adoption, capturing possible declines in human scholarly contribution. Another limitation is the uniform weighting of quality components ($\gamma = \alpha = \beta = \frac{1}{3}$) in the simulations, which, while practical, may not empirically reflect the relative importance of publications, citations, and effort across different fields. Field-specific weighting schemes, informed by expert surveys or citation network analyses, could address this imbalance.

The model's reliance on retraction data as a proxy for misconduct also poses challenges, as retractions are noisy indicators and many cases remain unreported, particularly in institutions with weak oversight. To improve robustness, alternative metrics such as post-publication peer review flags and preprint controversies can be incorporated in future work. Furthermore, the assumption of linear ethics decay ($\psi(D) = \omega D$) likely underestimates the risk of cascading failures, where trust in a journal or institution collapses suddenly. Nonlinear decay functions, such as sigmoidal or piecewise models, could better capture these tipping points. Finally, the conceptual subspaces representing originality, citation rigour, and other dimensions ($b_j$) lack operational definitions. These could be grounded in existing bibliometric indices, such as the OpenCitations Index of Crossref open DOI-to-DOI citations (COCI) for citation rigour or textual reuse detection for originality, to ensure empirical clarity.

The framework offers valuable practical implications and directions for refinement. One application is institutional benchmarking, where the ratio $\frac{AI_x(t)}{S}$ could serve as a public accountability metric, revealing gaps between declared policy tolerance and actual retraction rates. Institutions with disproportionately high ratios might undergo audits to align policy with practice. Early warning systems could also be developed by integrating the model with institutional data pipelines, such as plagiarism detection software and internal review reports, to flag departments where the probability of crossing the plagiarism tolerance threshold ($P$) exceeds safe levels. A dashboard visualising trends in academic damage metrics ($D_k(t)$) would help prioritise interventions.

## 5. Conclusions and Recommendations

This study explored the ethical implications and academic integrity risks associated with AI-mediated postgraduate supervision. Through a combination of theoretical modelling and empirical data analysis, several notable patterns emerged. While AI has the potential to greatly enhance research efficiency, accessibility, and analytical depth, the findings show that excessive or unguided use introduces significant risks. These include reduced originality in student work, heightened retraction vulnerability, and the erosion of meaningful mentorship. The study further demonstrates that the relationship between AI integration and research quality is non-linear, governed by diminishing marginal returns, reinforcing the need for deliberate and balanced adoption.

## 5.1 Institutional recommendations

To safeguard the integrity of postgraduate supervision in the era of AI, institutions must adopt a coherent and ethically grounded approach. This begins with the development of robust ethical AI frameworks that provide discipline-sensitive guidelines for appropriate AI use in academic work. Such frameworks should be reviewed regularly to reflect evolving technological and disciplinary realities (Papagiannidis, Mikalef, & Conboy, 2025). Alongside this, institutions should implement dynamic academic integrity systems capable of monitoring plagiarism tolerance thresholds and tracking key indicators of misconduct (including retraction rates) so that early intervention becomes both possible and consistent.

The study also highlights the importance of structured AI literacy training for both students and supervisors. This training should emphasise ethical AI use, citation conventions, and responsible academic writing practices to prevent overreliance on generative tools. Equally essential is the preservation of human-centred supervision. AI should complement, rather than replace, the interpersonal guidance, ethical mentorship, and scholarly dialogue that form the basis of quality postgraduate supervision (Koeszegi, 2024). Ensuring equitable access to AI tools and related training is another critical dimension; without such support, disparities in research quality may widen between students with differing levels of technological access.

## 5.2 Social Implications

The integration of AI into postgraduate supervision also carries important social implications. Equitable access to AI resources is vital to prevent the emergence of new academic divides that may disadvantage students from resource-constrained backgrounds. Moreover, the study underscores that human engagement remains central to fostering academic identity, confidence, and intellectual growth—elements that AI systems cannot replicate. Institutions should therefore prioritise supervisory models that balance technological assistance with sustained interpersonal mentorship, ensuring that social cohesion and inclusivity remain integral to the postgraduate experience. Furthermore, these findings challenge institutions to go beyond simple policy prohibitions and invest in pedagogical shifts that foster critical digital literacy among postgraduate students, moving them from mere functional use to sophisticated ethical reasoning. Ultimately, safeguarding the integrity of the supervisory process is paramount for maintaining public trust in the credentials awarded and ensuring the quality of the next generation of academic and industry research leaders.

## 5.3 Practical implications

From a practical perspective, institutions must adopt systematic mechanisms to monitor AI's long-term impact on research quality. This includes tracking how AI influences critical thinking, originality, and methodological rigour, and adjusting institutional policies as needed (Wits Centre for Learning, Teaching, and Development, 2024). Continuous evaluation enables institutions to detect emerging risks, refine training programmes, and ensure that AI remains a tool that enhances rather than undermines scholarly standards. A data-driven monitoring approach is, therefore, indispensable for maintaining quality assurance across postgraduate programmes. Specifically, the 'plagiarism tolerance threshold' identified through our econometric modelling offers a critical quantitative metric for policymakers, clearly signalling when reactive measures are no longer adequate and systemic change is immediately required. This necessitates the proactive development of transparent, adaptable institutional guidelines and the implementation of robust technological infrastructure capable of identifying and mitigating integrity risks before they escalate into widespread academic misconduct.

In conclusion, while AI offers powerful opportunities to enhance postgraduate research supervision, its integration must be accompanied by vigilant oversight, equitable support structures, and an unwavering commitment to academic integrity. By adopting ethical frameworks, strengthening mentorship, and instituting robust monitoring mechanisms, institutions can ensure that AI contributes positively to postgraduate education without compromising the values that underpin scholarly excellence.

## 6. Declarations

**Funding:** This research did not receive any external funding
**Conflicts of Interest:** The authors declare no conflicts of interest.
**Use of Artificial Intelligence:** The current work was supported with the assistance of artificial intelligence technologies (Grammarly) to assist with refining language for clarity, as confirmed by the authors.

# References

Acosta-Enriquez, B. G., Arbulu Ballesteros, M., Vilcapoma Pérez, C. R., Huamaní Jordan, O., Martin Vergara, J.A., Martel Acosta, R., Arbulu Perez Vargas, C. G., & Arbulú Castillo, J. C. (2025). AI in academia: How do social influence, self-efficacy, and integrity influence researchers' use of AI models? *Social Sciences & Humanities Open*, 11, https://doi.org/10.1016/j.ssaho.2025.101274

Ahmad, S. F., Han, H., Alam, M. M., Rehmat, M. K., Arraño-Muñoz, M., & Ariza-Montes, A. (2023). Impact of artificial intelligence on human loss in decision making, laziness and safety in education. *Humanities and Social Sciences Communications, 10*, 311. https://doi.org/10.1057/s41599-023-01787-8

Ali, O., Murray, P. A., Momin, M., Dwivedi, Y. K., & Malik, T. (2024). The effects of artificial intelligence applications in educational settings: Challenges and strategies. *Technological Forecasting and Social Change*, 199, 123076. https://doi.org/10.1016/j.techfore.2023.123076

Al-Jahwari, M., & Yousif, M. J. (2025). The impact of AI tools on education: ChatGPT in focus. *Artificial Intelligence & Robotics Development Journal, 4*(4), 314–336. https://doi.org/10.52098/airdj.20244430

Ali, Z. (2020). Artificial intelligence (AI): A review of its uses in language teaching and learning. *IOP Conference Series: Materials Science and Engineering, 769*(1), 12043. https://doi.org/10.1088/1757-899X/769/1/012043

Altmäe, S., Sola-Leyva, A., & Salumets, A. (2023). Artificial intelligence in scientific writing: A friend or a foe? *Reproductive BioMedicine Online, 47*(1), 3-9. https://doi.org/10.1016/j.rbmo.2023.04.009

Balk, B. M. (2024). Why is the Cobb-Douglas production function so popular? *Evolutionary and Institutional Economics Review, 21*(1), 1–20. https://doi.org/10.1007/s40844-024-00279-x

Bearman, M., Ryan, J., & Ajjawi, R. (2023). Discourses of artificial intelligence in higher education: A critical literature review. *Higher Education, 86*(2), 369–385. https://doi.org/10.1007/s10734-022-00937-2

Cobb, C. W., & Douglas, P. H. (1928). A theory of production. *The American Economic Review*, *18*(1), 139–165.

Crawford, J., Allen, K. A., Pani, B., & Cowling, M. (2024). When artificial intelligence substitutes humans in higher education: the cost of loneliness, student success, and retention. *Studies in Higher Education*, *49*(5), 883–897. https://doi.org/10.1080/03075079.2024.2326956

Kamalov, F., Santandreu Calonge, D., & Gurrib, I. (2023). New Era of Artificial Intelligence in Education: Towards a Sustainable Multifaceted Revolution. *Sustainability*, 15(16), 12451. https://doi.org/10.3390/su151612451

Koeszegi, S. T. (2024). AI @ work: Human empowerment or disempowerment?. In H. Werthner et al. (Eds.), *Introduction to digital humanism* (pp. 175–196). Springer. https://doi.org/10.1007/978-3-031-45304-5_12

Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S., Darwis, A., & Marzuki. (2023). Exploring artificial intelligence in academic essay: Higher education student's perspective. *International Journal of Educational Research Open, 5*, 100296. https://doi.org/10.1016/j.ijedro.2023.100296

Mauti, J. M. (2025). Ethical implications of artificial intelligence in university education. *East African Journal of Education Studies, 8*(1), 159-167. https://doi.org/10.37284/eajes.8.1.2583

Nguyen, M.-H., & Vuong, Q.-H. (2024). Artificial intelligence and retracted science. *AI & Society*. https://doi.org/10.1007/s00146-024-02090-z

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025). Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems, 34*(2), 101885. https://doi.org/10.1016/j.jsis.2024.101885

Wits Centre for Learning, Teaching, and Development. (2024). *Guidelines for GAI use in learning, teaching, and research*. University of the Witwatersrand. https://www.wits.ac.za/media/wits-university/learning-and-teaching/cltd/documents/Wits-CLTD-Guidelines-for-GAI-use-in-Learning-Teaching-Research-Dec2024.pdf

Yang, L., Liu, S., Tsoka, S., & Papageorgiou, L. G. (2016). Mathematical programming for piecewise linear regression analysis. *Expert Systems with Applications, 44*, 156–167. https://doi.org/10.1016/j.eswa.2015.08.034

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 1-27. https://doi.org/10.1186/s41239-019-0171-0

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*, 28. https://doi.org/10.1186/s40561-024-00316-7

# Integrating Artificial Intelligence in Postgraduate Supervision: Emergent Opportunities, Challenges, and Strategic Responses for Institutions

Obrain T. Murire[1]
Bramwell K. Gavaza[2]

AFFILIATIONS
[1]Faculty of Engineering, Built Environment & Information Technology, Walter Sisulu University, East London, South Africa.
[2]Faculty of Management & Public Administration Sciences, Walter Sisulu University, Butterworth, South Africa.

CORRESPONDENCE
Email: omurire@wsu.ac.za

REFERENCE
Murire, O. T., & Gavaza, B. K. (2026). Integrating artificial intelligence in postgraduate supervision: Emergent opportunities, challenges, and strategic responses for institutions. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 244–262). ERRCD Forum. https://doi.org/10.38140/obp4-2026-13



**Abstract:** The integration of AI (artificial intelligence) into postgraduate supervision has transformative potential for improving efficiency, communication, and research outcomes. This study explores both the opportunities and challenges associated with AI-based tools in postgraduate supervision. Using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, it systematically reviews peer-reviewed journal articles, conference papers, and reports to identify trends, benefits, and barriers to AI adoption. Thirty-seven studies were selected for final synthesis due to their direct relevance to the study's research questions. The findings indicate that AI offers significant opportunities, such as personalised learning paths, automated feedback cycles, enhanced research, increased visibility, and support for supervisor-student collaboration. Nonetheless, several major obstacles to widespread AI adoption remain. These include ethical implications, data privacy concerns, technical restrictions, and limitations in user digital literacy. Resistance from supervisors and students, driven by fears of AI supplanting human interaction, is also a notable barrier. The study proposes the development of regional and national guidelines for the phased implementation of AI tools in postgraduate supervision, starting with non-intrusive applications such as administrative support and progress tracking. Concurrently, training programmes for supervisors and students should be introduced to enhance digital literacy and cultivate a favourable attitude towards AI tools. Additionally, institutions must develop policies to address ethical issues and create a framework for human-AI collaboration that complements traditional supervision practices. This study advances scholarly understanding of AI's evolving role in postgraduate student supervision and provides actionable insights for institutions seeking to utilise AI effectively.

***Keywords:*** Artificial Intelligence; digital literacy; educational technology; higher education; postgraduate supervision; student engagement.

## 1. Introduction

Being an academic postgraduate research supervisor in a higher education institution is difficult and complicated (Wright, 2024). Efficient supervision in higher education, particularly at the doctoral level, is expected to meet high standards of quality, with strong quality assurance supporting all supervisory practices (Wright, 2024). However, the increasing importance of research in higher education may intensify the demands on research supervision, further burdening the already stretched student supervisors (Cowling et al., 2023).

Artificial Intelligence (AI) has been deemed beneficial for facilitating students' involvement during their postgraduate journey. AI technologies are transforming research and education by automating repetitive tasks, assisting with data analysis, and introducing innovative learning and assessment procedures (Zhai, Wibowo & Li, 2024). Wright (2024) highlights that the fast-evolving AI tools are significantly influencing the evaluation of academic work, especially in postgraduate studies. Li and Xing (2021) also underscore that AI enhances student participation and teamwork, empowering students to ask questions and engage in discussions asynchronously.

However, AI is considered a potential risk in the academic landscape, particularly concerning complications associated with the writing, submission, and supervision of postgraduate dissertations or theses. AI reshapes the dynamics, responsibilities, and trust involved in writing a dissertation or thesis (Dai et al., 2023). There is a risk that AI could undermine the fundamental objective of higher education, which is to foster critical thinking and intellectual development in students, ultimately diminishing the value of degrees (Cotton, Cotton & Shipway, 2023). While AI presents both opportunities and challenges in postgraduate supervision (Xia et al., 2024), how it will transform and influence higher education supervision practices remains unclear.

### 1.1 Overview of artificial intelligence and postgraduate supervision transformation

The use of AI in higher education has significantly transformed various academic processes, encompassing teaching, learning, administration, and increasingly, postgraduate supervision (Ahsan, Akbar & Kam, 2022; Murire, 2024). In light of rising postgraduate enrolments, limited supervisory capacity, and heightened expectations for quality research outputs, AI emerges as a transformative tool that has the potential to enhance the efficiency, accessibility, and effectiveness of supervision (Aladsani, 2025). Murire and Cilliers (2019) observe that there has been a global surge in postgraduate enrolments, driven by the knowledge economy's demand for highly skilled researchers and professionals. This phenomenon, commonly referred to as the massification of higher education, has exerted considerable pressure on the traditional supervision model (Manathunga, 2014; Murire & Cilliers, 2019). Moreover, the diversifying profile of postgraduate students, which includes part-time, distance, and international students, has further complicated supervisory processes (Holmes & Miao, 2023; George, 2023). Supervisors are required to navigate varying expectations, cultural diversity, and differing levels of research readiness among students. In such circumstances, maintaining regular communication, delivering timely and high-quality feedback, and effectively monitoring student progress becomes an ongoing challenge (Gavaza, 2024). Supervisors are expected to oversee larger cohorts of students, often across diverse disciplines, while concurrently addressing institutional research output targets, administrative duties, and teaching responsibilities (Day, 2023). Murire (2024) asserts that AI offers a viable solution to these evolving challenges.

Furthermore, Hopp and Speil (2021) contend that AI in postgraduate supervision is not intended to supplant human interaction but to complement and enrich the supervisory process. Consequently, AI technologies can assist with a variety of supervisory tasks as demands on supervisors continue to escalate.

### 1.2 Opportunities of AI in postgraduate supervision

One of AI's most valuable contributions to postgraduate supervision is its support for academic writing and the provision of timely, customised feedback (Smith et al., 2024). AI can also foster a more inclusive and equitable supervision atmosphere (Mannuru et al., 2023; Rodrigues et al., 2023). For example, AI-powered chatbots and virtual assistants can offer 24/7 assistance by answering repeated questions, explaining academic principles, or directing students to appropriate information (Hansen et al., 2025; Okoth, 2025). This is especially helpful for students who may be studying in different time zones or remotely (Titchener & Greene, 2023). Moreover, Day (2023) claims that AI-powered systems can facilitate asynchronous communication between supervisors and students, ensuring continuity in supervision despite resource constraints or geographic dispersion. These systems can record interactions, track milestones, and send reminders, thereby enhancing the management of the supervision process (Luckin et al., 2016).

Ahsan, Akbar, and Kam (2022) posit that supervisors generally face the dual challenges of providing scholarly guidance and managing time-consuming administrative tasks, such as monitoring progress, scheduling meetings, and overseeing submission deadlines (Hanim, Aripin & Lin, 2020). George (2023) states that AI can automate such tasks through intelligent dashboards and AI-driven project management software that send notifications, maintain records of milestones, and offer reports on student progress. Chan and Hu (2023) found that the use of AI in postgraduate supervision improves time management and allows supervisors to identify potential delays or bottlenecks early. Moreover, AI technologies like predictive analytics can examine student activity, engagement, and submission history to recognise threats such as disengagement or the likelihood of attrition. Early detection allows for timely action by supervisors and institutions, thereby strengthening postgraduate programme completion and retention rates (Williamson & Eynon, 2020).

George (2023) underscores that conducting a comprehensive literature review is often regarded as one of the most challenging and time-consuming stages of postgraduate research. Similarly, Okoth (2025) mentions that the literature review process requires not only the identification of relevant scholarly works but also the evaluation, synthesis, and organisation of these into a coherent narrative that justifies and supports the research focus. Murire and Cilliers (2019) argue that traditionally, this phase demands significant manual effort, critical thinking, and the ability to navigate vast volumes of academic data. However, Chauke et al. (2024) assert that advancements in AI are beginning to revolutionise this fundamental aspect of scholarly inquiry.

AI-powered tools such as Iris.ai, Research Rabbit, and Connected Papers provide researchers with robust capabilities for discovering, filtering, and organising relevant literature. These tools leverage techniques such as semantic similarity analysis, citation network mapping, and automated keyword extraction to go beyond basic keyword searches (Zawacki-Richter et al., 2019). For instance, by analysing conceptual linkages and citation patterns, these platforms can surface not only the most cited works but also hidden or emerging studies that are semantically related to a research query (Michel-Villarreal et al., 2023). This significantly enhances the discovery process and ensures a more robust and comprehensive foundation for academic writing (Lucey et al., 2021).

In addition to identifying pertinent studies, Dai et al. (2023) stated that these AI systems can visually map the evolution and interconnections of research topics over time. They can highlight emerging themes, predict future research directions, and filter out irrelevant or redundant sources, providing researchers with a more transparent and strategic understanding of their research landscape (Michel-Villarreal et al., 2023). This ability to map and navigate the intellectual terrain of a discipline enables students to make more informed decisions when framing their research questions, identifying gaps in existing knowledge, and positioning their studies within the broader academic conversation (Halse & Malfroy, 2010).

Supervisors and academic mentors also stand to benefit from these tools (Awdry, 2023). With access to real-time updates and analytics on new publications within their area of expertise, supervisors can remain current with the latest research developments, which can enhance the relevance and timeliness of the feedback they provide to their students (Dergaa et al., 2023). This continuous awareness contributes to more effective and evidence-informed supervision practices. However, there is a risk that overreliance on AI could deskill postgraduates and erode their critical thinking skills. Institutions should consider establishing AI ethics committees within postgraduate schools to monitor and guide the evolving impact of AI on academic supervision.

### 1.3 Challenges of AI in Postgraduate Supervision

A significant concern regarding the incorporation of artificial intelligence (AI) in postgraduate supervision is the preservation of academic integrity. Chauke et al. (2024) argue that the introduction of ChatGPT in higher education raises significant apprehensions about academic dishonesty. Ozguven et al. (2024) further assert that the integration of AI tools into higher education has led to both deliberate and inadvertent instances of manipulation by students in their utilisation of such technologies. The utilisation of AI for academic tasks in dishonest ways varies considerably. Dehouche (2021) also suggests that some students employ ChatGPT to fulfil assignments and conduct research, which could potentially undermine the integrity of their academic responsibilities.

The risk of dishonesty and cheating associated with the utilisation of AI spans various academic fields. Supervisors face the challenge of addressing issues related to academic cheating and dishonesty. Ahsan, Akbar, and Kam (2022) posit that there has been a notable increase in the number of scholarly articles addressing cheating in higher education in recent years, a trend that is partially driven by the growing accessibility of AI during assessments, as opposed to traditional in-person examinations, coupled with a rise in contract cheating. Beyond the concerns of academic integrity, some students have adopted AI-assisted writing as a substitute for their own work (Rodrigues et al., 2023). Hopp and Speil (2021) reported that approximately 22 per cent of students at an Austrian university confessed to engaging in plagiarism. Correspondingly, a study conducted by Chan and Hu (2023) involving students from six universities in Hong Kong highlighted issues related to accuracy and ethical considerations, particularly regarding plagiarism. AI can facilitate various forms of academic dishonesty, including data fabrication, plagiarism, examination cheating, improper collaboration, infringements of copyright, complicity in dishonest academic practices, and the alteration of bibliographic references, among other forms of academic misconduct (Rodrigues et al., 2023). The serious issue of plagiarism and the consequent disciplinary actions connected to AI continues to persist (Wright, 2024). The emergence of AI exacerbates the risks associated with academic cheating by potentially making academic services accessible to a broader cohort of students who may not perceive AI utilisation as dishonest or who previously could not afford essay mill services (Cotton et al., 2023). In such circumstances, Cotton et al. (2023) contend that university supervisors are compelled to meticulously consider the structure of their assessments and procedures to effectively communicate the issues of academic dishonesty and cheating to students while minimising their occurrence.

Even with the numerous advantages associated with certain AI tools, Day (2023) cautions that their utilisation in research may adversely affect students' academic achievements. The outputs generated by AI tools might include fabricated citations, references, and responses, which pose significant threats to the integrity of academic work. Holmes and Miao (2023) further affirm that AI-generated answers can be based on fictitious references, potentially undermining the technology's credibility as a reliable resource for research and education. George (2023) suggests that AI models can propagate misconceptions if deployed in academia without rigorous validation and effective supervision. Aladsani (2025) also posits that some students, having encountered erroneous references produced by AI, have come to realise the necessity of conducting thorough reference checks to avoid sanctions from their supervisors.

A major concern regarding the use of AI in the academic landscape, particularly regarding the supervision of postgraduate students, is the erosion of critical thinking, analysis, and writing skills.Ateeq et al. (2024) argue that AI tools designed to assist research and provide expedited solutions can easily be misused. Students may rely excessively on AI resources, which could impair their analytical and problem-solving abilities (Okoth, 2025). Dergaa et al. (2023) also emphasise that AI has the potential to diminish vital cognitive skills, such as critical and analytical thinking, as well as decision-making. An overreliance on AI tools for feedback and guidance may disrupt direct communication between students and their supervisors, thereby hindering the development of a robust mentoring relationship and the sharing of implicit knowledge. Seo et al. (2021) contend that the interaction between students and supervisors constitutes the most critical form of academic interaction and that AI diminishes this engagement. Nonetheless, AI cannot wholly replace the guidance and emotional support that human supervisors provide to postgraduate students (Cowling et al., 2023). AI lacks the specialised knowledge that skilled supervisors and professors possess (Michel-Villarreal et al., 2023). Although AI platforms are typically effective in delivering comprehensive strategic responses and aggregating broad information, they struggle to provide in-depth insights regarding organisations, individuals, and locations that have received diminished global attention (Cowling et al., 2023).

The use of artificial intelligence (AI) may engender unfair assessments during postgraduate supervision. Cotton et al. (2023) articulate that the application of AI encompasses the generation of high-quality written research reports, implying that those who possess access to AI tools might hold an unfair advantage over their peers who do not. Holmes and Miao (2023) express apprehensions regarding the influence wielded by AI developers, which may lead to the potential for bias in information accessibility and the marginalisation of particular perspectives. The biases inherent in the data can inadvertently persist through the outcomes of AI, thereby affecting both accuracy and impartiality (Aladsani, 2025). Furthermore, Dai et al. (2023) highlight that the inherent bias within AI algorithms, along with concerns pertaining to fairness and the risk of misuse, constitutes fundamental issues that necessitate resolution by supervisors. Should the data employed for training be skewed, the AI may yield biased results or recommendations, which could adversely impact specific students or research domains. This scenario may culminate in inequities during the assessment processes of postgraduate supervision (Cotton et al., 2023).

Certain supervisors exhibit a lack of comprehensive understanding regarding AI tools, as there exist limited policies or established strategies governing their usage, whether to support, oppose, accept, or prohibit them (Wright, 2024). Guo and Wang (2023) assert that AI is prompting postgraduate supervisors to modify assessments to address the challenges associated with its utilisation. Ramberg and Modin (2019) further contend that academic dishonesty demonstrates an inverse relationship with the understanding and endorsement of academic integrity policies. George (2023) argues that AI may produce stereotypical, offensive, or prejudiced content in the absence of adequate supervision if employed in postgraduate education.

### 1.4 Problem statement

Postgraduate students increasingly operate within complex and specialised academic fields that necessitate extensive supervisory guidance, advanced research skills, and continuous academic support (Dai et al., 2023). However, many higher education institutions encounter persistent supervisory challenges, including limited supervisor availability, administrative burdens, and difficulties in providing timely, individualised, and discipline-specific feedback (Halse & Malfroy, 2010; Igumbor et al., 2022). These pressures are further exacerbated by rising postgraduate enrolments and shortages of qualified PhD-level supervisors in certain departments (Mahlangu, 2021). Although institutions have implemented technology-enhanced systems and hybrid supervision models to alleviate these challenges (Wisker et al., 2021), traditional supervision structures remain under strain. Emerging developments in Artificial Intelligence (AI) present promising opportunities to enhance supervisory efficiency, support scholarly writing, improve literature review processes, and provide personalised feedback (Chu et al., 2022). However, despite its potential, the integration of AI raises significant concerns regarding academic integrity, ethical use, student dependence, cognitive development, and algorithmic bias. What remains insufficiently understood is how AI can be responsibly leveraged to support postgraduate supervision while mitigating these risks. Accordingly, there is a need to investigate the opportunities that AI presents for strengthening supervision, the challenges and ethical issues it introduces, and the strategies required to ensure its responsible and effective adoption. This chapter addresses these gaps by examining the opportunities, challenges, and institutional enablers associated with the integration of AI into postgraduate supervision. Based on this, the following questions were raised:

- What are the opportunities associated with the integration of AI in postgraduate supervision?
- What are the challenges associated with the integration of AI in postgraduate supervision?
- Which strategies should institutions use to overcome the resistance to AI adoption in postgraduate supervision?

## 2. **Methodology**

This chapter adopts a qualitative research design to explore both the opportunities and challenges associated with AI-based tools in postgraduate supervision. A qualitative research approach is particularly suitable for this study because it enables a rich, detailed understanding of complex processes, such as supervision practice and technology integration, from the insiders' perspectives (Snyder, 2019).

Data collection was conducted using a systematic literature review (SLR). The SLR served as the conceptual framework and was designed to synthesise peer-reviewed journal articles, conference articles, and books published between 2015 and 2024. The literature search focused on relevant topics such as AI in higher education, writing aids in academia, research management tools, and ethical considerations of employing AI in supervision. These were searched through scholarly databases like Scopus, Web of Science, Google Scholar, and ERIC using keywords such as “Artificial Intelligence AND postgraduate supervision,” “AI tools in research,” and “AI in academic mentoring.” The selection criteria prioritised recency, relevance, and peer-reviewed literature. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were adopted in the review process, highlighting transparency and methodological precision (Moher et al., 2009).

The selection and review of literature were guided by the PRISMA guidelines (Moher et al., 2009), which provided a structured and transparent review process. The initial database search yielded 151 records. After removing duplicates, 91 articles remained. Screening based on titles and abstracts, along with the application of predefined inclusion and exclusion criteria, resulted in 71 articles for full-text review. From this pool, 37 articles included in the reference list were selected for final synthesis due to their direct relevance to the study's research questions.

Two authors collaboratively conducted the study, each bringing complementary strengths to the research process. Both authors led the systematic review process, including database searches, literature screening, and article selection. They were equally involved in conceptualising the research design, defining the scope, and developing the framework for analysis. The writing process was collaborative and iterative, with drafts shared via cloud-based platforms to enable joint editing, critical reflection, and refinement. Regular discussions were held to ensure alignment and maintain research integrity. Additionally, both authors independently applied the CASP checklist and engaged in peer debriefing to enhance the rigour and trustworthiness of the study.

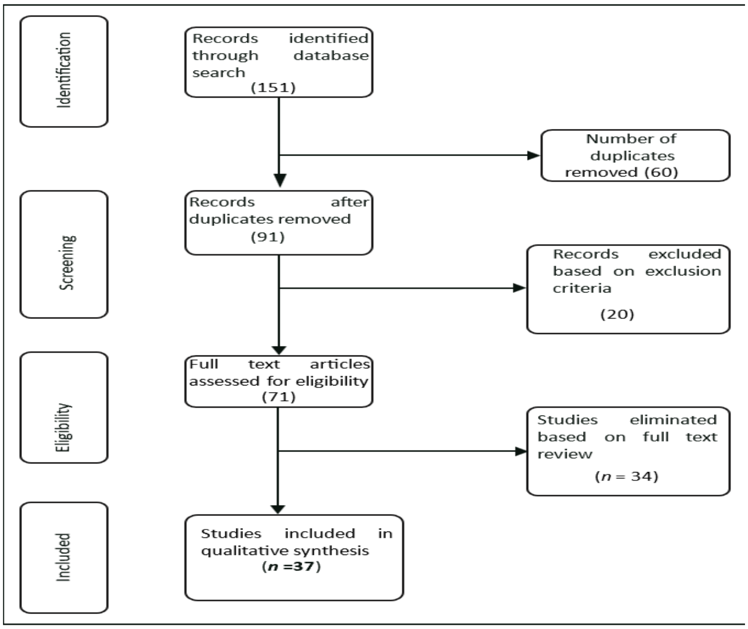



***Figure 3:*** *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*

## 2.1 Data extraction and synthesis

While the PRISMA method offered rigour in identifying relevant literature, it constrained the study's capacity to explore emergent, practice-based themes that a grounded theory approach might have captured. Additionally, the absence of a longitudinal design similarly limits our ability to detect the evolving impacts of AI over time. Given the rapid nature of advancements in AI technologies, some findings may become outdated quickly. Future studies should involve larger, more diverse samples and employ mixed methods to enhance generalisability and provide comparative findings. Longitudinal studies would facilitate tracking the continued impact of AI on postgraduate supervision, and discipline-specific studies might uncover the adoption and experience of AI across various academic fields. Further research into the ethical, psychological, and institutional dimensions of AI adoption is also required to enable responsible and context-sensitive use.

## 3. Findings and Discussion

This section presents the findings thematically, guided directly by the study's three research questions. The new structure enhances conceptual flow, reflects the logical progression of the study objectives, and intentionally foregrounds the authors' analytical voice. Rather than merely summarising previous studies, the discussion critically interprets what these findings mean for postgraduate supervision practice and policy in the evolving AI landscape.

### 3.1 Opportunities associated with integrating AI into postgraduate supervision

Across the reviewed literature, one of the strongest themes is the transformative potential of AI to enhance the quality, efficiency, and accessibility of postgraduate supervision. Our analysis shows that AI introduces significant improvements in three key supervisory domains: administrative efficiency, academic writing and research support, and student engagement.

Firstly, AI tools substantially reduce supervisors' routine administrative workload. Applications such as Grammarly, Turnitin, and AI-enabled dashboards automate substantial portions of text analysis, error detection, progress tracking, and time management (Awdry, 2023; Zawacki-Richter et al., 2019). As authors, we view this as a crucial step towards repositioning supervisors to focus on higher-order, conceptual guidance, which has long been identified as lacking in massified higher education environments (Murire & Cilliers, 2019). By alleviating supervisors of administrative burdens, AI creates opportunities for deeper intellectual engagement with postgraduate students. Secondly, AI significantly enhances students' academic writing and research capabilities. Tools such as Iris.ai, Connected Papers, and Research Rabbit provide sophisticated literature discovery, semantic search, and citation network mapping (Chu et al., 2022; Dergaa et al., 2023). In our interpretation, these tools not only accelerate the literature review process but also democratise access to advanced research support, particularly for students who may lack extensive disciplinary familiarity. While earlier research characterised the literature review as one of the most cognitively demanding stages of postgraduate research, our findings demonstrate that AI has begun to reshape this landscape by making scholarly inquiry more navigable, transparent, and efficient.

Thirdly, AI enhances student engagement, accessibility, and personalised learning. Chatbots, intelligent feedback systems, and translators offer 24/7 academic support, enabling remote, part-time, and international students to access timely guidance (Williamson & Eynon, 2020; Smutny & Schreiberova, 2020). Our analysis further indicates that adaptive learning systems foster autonomy, motivation, and confidence among postgraduate students—an essential ingredient for successful research trajectories (Holmes et al., 2021; Taylor, 2019). Collectively, these findings illustrate that AI is not merely a tool for efficiency; it represents a strategic enhancement to supervision ecosystems. When used responsibly, AI has the potential to equalise access, elevate academic writing skills, and reduce systemic supervisory bottlenecks that have historically undermined postgraduate success in resource-constrained institutions.

### 3.2 Challenges associated with integrating AI into postgraduate supervision

Despite the opportunities presented by artificial intelligence (AI), the literature highlights significant challenges that necessitate careful institutional oversight. Our thematic analysis identifies four primary concerns: academic integrity, data reliability, erosion of cognitive skills, and algorithmic bias and inequity.

The first and most prominent challenge pertains to academic integrity. Numerous studies have documented an increase in plagiarism, contract cheating, and the misuse of AI tools for generating assignments and research content (Cotton et al., 2023; Dehouche, 2021; Rodrigues et al., 2023). We interpret this trend as indicative of inadequate regulatory frameworks and inconsistent digital literacy among student populations. The growing sophistication of generative AI tools has blurred traditional boundaries of authorship, rendering academic misconduct both easier to perpetrate and more difficult to detect. Supervisors are faced with the dual burden of policing misconduct while also guiding the ethical use of AI. The second theme concerns the reliability and credibility of AI-generated outputs. Findings highlight the occurrence of fabricated citations, hallucinated references, and inaccurate responses produced by AI tools (Day, 2023; Holmes & Miao, 2023). These errors not only compromise students' research integrity but also shift additional verification responsibilities onto supervisors. From our perspective, this underscores the necessity for supervisors to remain epistemic gatekeepers—ensuring that AI complements rather than compromises academic standards. Thirdly, extensive use of AI may erode critical thinking and analytical reasoning skills. Studies suggest that an overreliance on AI-driven text generation and feedback diminishes students' cognitive engagement with content (Okoth, 2025; Dergaa et al., 2023). We contend that while AI can scaffold learning, excessive dependence risks producing graduates who lack independent scholarly judgment. This challenge underscores the importance of supervisory intervention in maintaining reflective, inquiry-driven learning.

Fourthly, AI introduces risks of algorithmic bias and unequal learning experiences. AI systems trained on skewed datasets may produce biased recommendations or marginalise certain disciplinary or cultural perspectives (Aladsani, 2025; Dai et al., 2023). Furthermore, variations in digital literacy mean that students with greater technological familiarity disproportionately benefit. Our analysis suggests that, in the absence of institutional safeguards, the adoption of AI could inadvertently exacerbate inequities in postgraduate education rather than mitigate them. In summary, while AI provides powerful support mechanisms, its associated risks are substantial. Effective integration necessitates robust governance structures, enhanced academic integrity policies, and a shift in supervisory practices towards ethical, critical, and reflective engagement with AI tools.

### 3.3 Strategies to overcome resistance and maximise responsible AI adoption

The literature demonstrates that resistance to AI adoption is partly rooted in institutional uncertainty and partly in psychological and ethical discomfort among both supervisors and students. To address these barriers, our synthesis identifies three strategic imperatives: capacity building, policy development, and human-centred supervision models.

First, digital literacy training emerges as a critical enabler. Institutions must develop structured programmes that train supervisors and students in the ethical use of AI, the verification of AI outputs, and the critical appraisal of AI-generated information (Kelly, 2023; UNESCO, 2021). We contend that digital literacy should not be optional; it must become a core component of postgraduate training to ensure equitable access and responsible adoption. Second, clear and enforceable AI governance policies are needed. The literature emphasises gaps in institutional protocols addressing the use of generative AI, ethical boundaries, plagiarism, and data privacy (Wright, 2024; Holmes & Miao, 2023). We argue that institutions must move beyond generic regulatory statements and develop discipline-specific guidelines co-created with supervisors and postgraduate students. Such policies will provide clarity on permissible AI support and delineate academic misconduct more explicitly.

Third, AI integration must be grounded in human-centred supervision. Our analysis confirms that supervisors remain irreplaceable for emotional support, intellectual mentorship, and disciplinary expertise—elements that AI cannot replicate (Cowling et al., 2023). Therefore, responsible AI adoption should complement, rather than replace, the supervisory relationship. Supervisors should use AI strategically for feedback generation, progress monitoring, and administrative assistance while maintaining human-led guidance for conceptualisation, research logic, and doctoral identity formation. We conclude that responsible AI adoption hinges not on technological capability but on institutional readiness, pedagogical integrity, and sustained human involvement. AI can enhance postgraduate supervision only when supervisors and institutions engage with it critically, reflexively, and ethically.

## 4. Conclusions

While AI offers transformative potential for postgraduate supervision, its adoption must proceed with a critical, human-centred perspective. Technology alone cannot address the relational, ethical, and pedagogical complexities inherent in research supervision. Thus, the future of supervision lies not in replacing the human mentor but in reimagining supervisory practices wherein AI becomes an ethical, equitable, and transparent partner in fostering research excellence. The integration of AI into postgraduate supervision represents a fundamental paradigmatic shift in the manner in which universities support research, mentorship, and academic growth. As explored in the course of this chapter, AI presents exceptional opportunities to enhance the quality, productivity, and accessibility of supervision through tools that facilitate literature reviews, provide writing assistance, offer feedback mechanisms, and generate research metrics. These tools can minimise administrative burdens for supervisors, ensure timely feedback, and enable students to evolve into more autonomous and reflective researchers. Simultaneously, the study has illuminated a range of challenges that must be navigated with care. These include ethical concerns, data privacy risks, potential overreliance on AI tools, and disparities in digital literacy among students and supervisors.

Moreover, the findings underscore a clear need for human-centred frameworks that preserve the interpersonal, developmental, and dialogical aspects of postgraduate supervision while leveraging AI's capabilities to augment, rather than replace, these core academic relationships. To ensure that AI enhances rather than undermines the excellence of postgraduate education, institutions must actively formulate policies, build digital literacy through targeted training programmes, and cultivate a culture of critical engagement with AI technologies. A reflective, ethically driven approach to AI integration, based on collaboration, transparency, and academic integrity, will be key to shaping the future of postgraduate supervision in a manner that benefits both students and supervisors. Lastly, AI should not be perceived as a threat to the traditional supervisory framework but rather as a powerful adjunct that, if ethically leveraged, can transform the postgraduate research process into a more open, efficient, and creative endeavour. As the higher education landscape continues to evolve, embracing this transformation with both openness and caution will be crucial in preparing future scholars for success in a digitally enhanced world.

## 5. Recommendations

Several practical recommendations can be made to support the effective and ethical adoption of AI in postgraduate supervision practices.

In developing these recommendations, we consciously emphasise a balanced perspective. Drawing from our review and critical reflection, it is evident that without intentional, structured digital literacy training, there is a risk that AI could exacerbate inequalities in postgraduate supervision, privileging digitally fluent students and supervisors while marginalising others. We therefore advocate for institutional policies that are not merely compliance-based but are dialogic and co-created with end-users to ensure sustainable, ethical AI integration in supervision contexts.

The incorporation of AI must foster critical thinking among postgraduate students. Universities that incorporate AI ethics into their programmes should explicitly address AI plagiarism to encourage students to develop original, critical, and creative thinking skills. Supervisors ought to evaluate AI system outputs, rectify inaccuracies, provide context, and cultivate strong reasoning abilities in students. Educational institutions should highlight the importance of supervisors and teaching personnel in guiding students, confirming the reliability of AI-generated information, and offering additional context when necessary. Supervisors should focus on structuring courses to ensure that AI promotes problem-solving, critical thinking, and creativity. They should also integrate AI technologies into academic courses, developing assessments or assignments that permit students to practise and enhance their writing abilities through the utilisation of AI.

AI is intended to assist in providing guidance for postgraduate students. Supervisors should design courses in a manner that accommodates the use of AI in an engaging style. There is a need for an integrated system that incorporates AI, which can be utilised to deliver mini-lectures, tests, and explanations tailored to each student, thereby enhancing outcomes in postgraduate courses. This process requires supervisors to supply instructional prompts and rubrics for educational objectives, scope, and format requirements. Additionally, academic institutions must offer materials and training to familiarise students with AI technologies and encourage the ethical utilisation of such tools in academic writing.

There is a necessity to restructure policies to ensure that appropriate regulations govern AI-supported postgraduate student supervision. Higher education institutions should reconsider their policies, curricula, and teaching procedures to better prepare students for a future in which AI technologies will be predominant. We recommend that postgraduate degree policies at the institutional level incorporate university protocols on the ethical usage of generative AI as one component of these guidelines and provide more detailed recommendations than those offered by external regulatory agencies. Academic institutions must develop and revise curricula and policies to address the changes brought about by AI; they must adapt to evolving educational requirements and ensure equitable access to AI.

Reflecting on the study, it became evident that the success of AI integration depends not solely on technological innovation but also on cultivating a culture of trust and openness between supervisors and students.

### 6.1 Social and practical implications of the study

The findings of this chapter carry significant social and practical implications for higher education institutions, supervisors, postgraduate students, and policymakers. From a social perspective, the integration of artificial intelligence (AI) into postgraduate supervision has the potential to mitigate long-standing inequalities in access to academic support. Students who study part-time, remotely, or in under-resourced settings stand to benefit from AI-enabled tools that provide continuous feedback, translation support, accessibility features, and personalised academic guidance. This democratisation of support can bridge gaps between students with varying levels of research readiness, digital literacy, and access to supervision. Concurrently, the study highlights the need for institutions to address emerging risks such as academic dishonesty, unequal access to digital tools, and reliance on AI-generated content, which can exacerbate existing inequities if not carefully managed.

Practically, the study demonstrates that AI holds substantial potential to reduce supervisory workloads, streamline administrative processes, and enhance the efficiency of supervision. Supervisors may allocate more time to intellectual mentorship, an aspect of postgraduate education that AI cannot replace, while relying on AI to manage routine tasks such as organising literature, providing writing feedback, tracking progress, and detecting plagiarism. For institutions confronting high student-to-supervisor ratios and shortages of qualified supervisors, AI presents a viable solution for enhancing the overall quality of supervision. However, the practical implications also underscore the urgent need for comprehensive institutional policies governing the use of AI. Universities must establish clear guidelines on the ethical use of AI, academic integrity, and accountability, alongside structured training programmes to build digital literacy among supervisors and students.

Overall, the study suggests that successful AI adoption in postgraduate supervision requires a balanced, human-centred strategy that preserves the relational and developmental aspects of supervision while enhancing them through technological support. Institutions that invest in ethical frameworks, training initiatives, and equitable access to AI tools can leverage AI as a catalyst for improved research quality, enhanced student experience, and sustainable supervision models across diverse academic contexts.

## 7. Limitations of the study

Although the PRISMA method provided a rigorous and systematic approach to identifying relevant literature, it also imposed certain constraints. Its structured nature limited the study's capacity to capture emergent, practice-based themes that might have been uncovered through more interpretive methodologies, such as grounded theory. Furthermore, the study did not adopt a longitudinal design, which restricts insights into how the impacts of AI evolve over time. Given the rapid pace of technological advancement, some findings may quickly become outdated.

The scope of the reviewed studies was also constrained by the size and diversity of available samples, which affects the generalisability of the conclusions. Future research should consider larger and more heterogeneous participant groups, as well as mixed methods approaches, to provide richer and more comparative perspectives.

Longitudinal research is essential for tracking the ongoing influence of AI on postgraduate supervision, while discipline-specific investigations may reveal variations in AI adoption and experiences across academic fields. Furthermore, a deeper examination of the ethical, psychological, and institutional implications of AI use is necessary to facilitate the responsible and context-sensitive integration of these technologies in higher education.

## 8. Declarations

**Funding:** This research did not receive any external funding

**Conflicts of Interest:** The authors declare no conflict of interest.

**Use of Artificial Intelligence:** The current work was created without the assistance of artificial intelligence technologies.

# References

Awdry, R. (2023). *A12 making friends with AI*. N-TUTORR Academic Integrity Masterclass, 23 March. Available at: https://vimeo.com/815310897

Ahsan, K., Akbar, S., & Kam, B. (2022). Contract cheating in higher education: A systematic literature review and future research agenda. *Assessment & Evaluation in Higher Education, 47*(4), 523–539. https://doi.org/10.1080/02602938.2021.1931660

Aladsani, H. K. (2025). Developing postgraduate students' competencies in generative artificial intelligence for ethical integration into academic practices: A participatory action research. *Interactive Learning Environments*, 1–19. https://doi.org/10.1080/10494820.2025.2487517

Ateeq, A., Alzoraiki, M., Milhem, M., & Ateeq, R. A. (2024). Artificial intelligence in education: Implications for academic integrity and the shift toward holistic assessment. *Frontiers, 9*, 1-14. https://doi.org/10.3389/feduc.2024.1470979

Bailey, D. R., & Zawacki-Richter, O. (2021). *Artificial intelligence in higher education: Applications and implications*. Springer.

Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O., & Kerres, M. (2021). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *International Journal of Educational Technology in Higher Education, 18*(1), 1–30. https://doi.org/10.1186/s41239-021-00249-7

Chan, C. K. Y. (2023). Is AI changing the rules of academic misconduct? An in-depth look at students' perceptions of 'AI-giarism'. *arXiv preprint arXiv:2306.03358*.

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*(1), 1-18. https://doi.org/10.1186/s41239-023-00411-8

Chauke, T. A., Mkhize, T. R., Methi, L., & Dlamini, N. (2024). Postgraduate students' perceptions on the benefits associated with artificial intelligence tools for academic success: The use of the ChatGPT AI tool. *Journal of Curriculum Studies Research, 6*(1), 44-59. https://doi.org/10.46303/jcsr.2024.4

Chu, H. C., Hwang, G. H., Tu, Y. F., & Yang, K. H. (2022). Roles and research trends of artificial intelligence in higher education: A systematic review of the top 50 most-cited articles. *Australasian Journal of Educational Technology, 38*(3), 22–42. https://doi.org/10.14742/ajet.7526

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Cowling, M., Crawford, J., Allen, K. A., & Wehmeyer, M. (2023). Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research supervision. *Australasian Journal of Educational Technology, 39*(4), 89–103.

Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology, 39*(4), 74–88. https://doi.org/10.14742/ajet.8843

Day, T. (2023). A preliminary investigation of fake peer-reviewed citations and references generated by ChatGPT. *The Professional Geographer, 75*(6), 1024–1027. https://doi.org/10.1080/00330124.2023.2190373

Dehouche, N. (2021). Plagiarism in the age of massive generative pre-trained transformers (GPT-3). *Ethics in Science and Environmental Politics, 2*, 17–23. https://doi.org/10.3354/esep00195

Dergaa, I., Chamari, K., Zmijewski, P., & Saad, H. B. (2023). From human writing to artificial intelligence generated text: Examining the prospects and potential threats of ChatGPT in academic writing. *Biology of Sport, 40*(2), 615–622. https://doi.org/10.5114/biolsport.2023.125623

Ellison, A. M., & Patel, M. V. (2022). *Success in navigating your student research experience: Moving forward in STEMM*. Springer Nature.

Espinoza Vidaurre, S. M., Velasquez Rodriguez, N. C., Gambetta Quelopana, R. L., Martinez Valdivia, A. N., Leo Rossi, E. A., & Nolasco-Mamani, M. A. (2024). Perceptions of artificial intelligence and its impact on academic integrity among university students in Peru and Chile: An approach to sustainable education. *Sustainability, 16*(20), 1–13.

Gavaza, B. K. (2024). *The impact of Covid-19 on teaching and learning in a historically disadvantaged university in Eastern Cape Province* [Master's dissertation, Walter Sisulu University].

George, A. S. (2023). The potential of generative AI to reform graduate education. *Partners Universal International Research Journal, 2*(4), 36–50.

Guo, K., & Wang, D. (2023). To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing. *Education and Information Technologies,* 12(3), 1-23. https://doi.org/10.1007/s10639-023-12146-0

Halse, C., & Malfroy, J. (2010). Retheorizing doctoral supervision as professional work. *Studies in Higher Education, 35*(1), 79–92. https://doi.org/10.1080/03075070902906798

Hanim, N., Aripin, N., & Lin, N. M. (2020). Exploring the connection between critical thinking skills and academic writing. *International Journal of Asian Social Science, 10*(2), 118–128. https://doi.org/10.18488/journal.1.2020.102.118.128

Hansen, R. R., Hougaard, R. F., Moller, K. L., Lindberg, A. B., Nielsen, T. A., & Prilop, C. N. (2025). The effects of an AI feedback coach on students' peer feedback quality, composition, and feedback experience. *Læring og Medier, 31*, 1–31.

Holmes, W., & Miao, F. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing.

Hopp, C., & Speil, A. (2021). How prevalent is plagiarism among college students? Anonymity preserving evidence from Austrian undergraduates. *Accountability in Research, 28*(3), 133–148. https://doi.org/10.1080/08989621.2020.1804880

Kelly, B. (2023). *Generative artificial intelligence guidelines for educators*. NAIN.
Igumbor, J. O., Bosire, E. N., Karimi, F., Katahoire, A., Allison, J., Muula, A. S., Peixoto, A., Otwombe, K., Gitau, E., Bondjers, G., Fonn, S., & Ajuwon, A. (2022). Effective supervision of doctoral students in public and population health in Africa: CARTA supervisors' experiences, challenges and perceived opportunities. *Global Public Health, 17*(4), 496–511. https://doi.org/10.1080/17441692.2020.1864752
Li, C., & Xing, W. (2021). Natural language generation using deep learning to support MOOC learners. *International Journal of Artificial Intelligence in Education, 31*(2), 186–214. https://doi.org/10.1007/s40593-020-00235-x
Li, M., Gibbons, J., Meng, H., & Taha, G. (2023). Graduate students' experience on using ChatGPT in education: A narrative inquiry. *SSRN*. https://doi.org/10.2139/ssrn.4452108
Lucey, B., Spriggs, R., O'Brien, K., Murphy, C., & Wright, A. (2021). The mentored mentoring approach to lecturer-and peer-led inter-institutional instruction among undergraduate biomedical science students: A case study. *AISHE-J*, 13(3), 1–21. https://doi.org/10.21427/xxx
Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education.
Mahlangu, V. P. (2021). Exploring challenges of supervising postgraduate students in open distance learning in higher education settings. *Bulgarian Comparative Education Society*, 19, 119–125. https://doi.org/10.3390/xxx
Manathunga, C. (2014). *Intercultural postgraduate supervision: Reimagining time, place and knowledge.* Routledge.
Mannuru, N. R., Shahriar, S., Teel, Z. A., Wang, T., Lund, B. D., Tijani, S., Pohboon, C. O., Agbaji, D. A., Alhassan, J., Galley, J., Kousari, R., Ogbadu-Oladapo, L., Saurav, S. K., Srivastava, A., Tummuru, S. P., Uppala, S., & Vaidya, P. (2023). Artificial intelligence in developing countries: The impact of generative artificial intelligence (AI) technologies for development. *Information Development*, 02666669231200628. https://doi.org/10.1177/02666669231200628
Marr, B. (2023, May 19). A short history of ChatGPT: How we got to where we are today. *Forbes*. https://www.forbes.com/sites/bernardmarr/2023/05/19/a-short-history-of-chatgpt-how-we-got-to-where-we-are-today/
Murire, O., & Cilliers, L. (2019). Critical success factors to improve the adoption of social media in teaching and learning: A case study at a traditional university. *International Journal of Interactive Mobile Technologies (iJIM)*, *13*(03), 81–94.https://doi.org/10.3991/ijim.v13i03.8485
Murire, O. T. (2024). Artificial intelligence and its role in shaping organisational work practices and culture. *Administrative Sciences, 14*(12), 316. https://doi.org/10.3390/admsci14120316
Okoth, M.O. (2025, February). Use of AI in research and postgraduate supervision: Balancing efficiency and academic integrity. *Scholar Media Africa*. https://scholarmedia.africa/education/higher-ed-bar/use-of-ai-in-research-and-postgraduate-supervision-balancing-efficiency-and-academic-integrity/#:~:text=While%20AI%20offers%20significant%20advantages,generating%20text%2C%20and%20analyzing%20data
Ozguven, M., Vahed, A., Akhal, K., & Garcia, A. B. (2024). Preserving academic integrity in AI-generated assessments: A case study in entrepreneurship at a Sino-foreign university.

*African Journal of Inter/Multidisciplinary Studies, 6*(1), 1-11. https://doi.org/10.51415/ajims.v6i1.1581

Rababah, L. M., Rababah, M. A., & Al-Khawaldeh, N. N. (2024). Graduate students' ChatGPT experience and perspectives during thesis writing. *International Journal of Engineering Pedagogy, 14*(3).

Ramberg, J., & Modin, B. (2019). School effectiveness and student cheating: Do students' grades and moral standards matter for this relationship? *Social Psychology of Education, 22*, 517–538.

Rodrigues, M., Silva, R., Borges, A. P., Franco, M., & Oliveira, C. (2023). Artificial intelligence: Threat or asset to academic integrity? A bibliometric analysis. *Kybernetes, 54*(5), 2939-2970. https://doi.org/10.1108/K-09-2023-1666

Seo, K., Tang, J., Roll, I., Fels, S., & Yoon, D. (2021). The impact of artificial intelligence on learner–instructor interaction in online learning. *International Journal of Educational Technology in Higher Education, 18*, 1-23.

Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. *Polity Press.*

Smith, S., Tate, M., Freeman, K., Walsh, A., Ballsun-Stanton, B., Hooper, M., & Lane, M. (2024). A university framework for the responsible use of generative AI in research. *arXiv preprint arXiv:2404.19244*.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, *104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039

Taylor, S. (2019). *A framework for good supervisory practice*. UK: Council for Graduate Education. https://doi.org/10.13140/RG.2.2.12696.34564

Titchener, H., & Greene, B. (2023). Building research integrity. In S. Gasson, J. Blacker, I. Stoodley, A. Winter, & C. Bruce (Eds.), *Confident supervisors: Creating independent researchers*. Available at: https://jcu.pressbooks.pub/confidentsupervisors (Accessed: 20 April 2025).

UNESCO. (2021). *Ethics of artificial intelligence: UNESCO recommendation*. https://unesdoc.unesco.org/ark:/48223/pf0000381137

Williamson, B., & Eynon, R. (2020). Historical threads, missing links, and future directions in AI in education. *Learning, Media and Technology*, *45*(3), 223–235. https://doi.org/10.1080/17439884.2020.1798995

Wisker, G., McGinn, M. K., Bengtsen, S. S., Lokhtina, I., He, F., Cornér, S., Leshem, S., Inouye, K., & Lofstrom, E. (2021). Remote doctoral supervision experiences: Challenges and affordances. *Innovations in Education and Teaching International*, *58*(6), 612–623. https://doi.org/10.1080/14703297.2021.1991427

Wright, A. (2024). Postgraduate supervision in a ChatGPT world: What's next? In *10th International Conference on Higher Education Advances (HEAd'24)*. Valencia, 18-21 June 2024. https://doi.org/10.4995/HEAd24.2024.17244

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21(1), 1-14.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments*, 11(1), 1-8. https://doi.org/10.1186/s40561-024-00316-7

# Effective AI Integration in Postgraduate Supervision Practices: Policy Implications for South Africa

Petronella Elize Saal[1]
Krish Chetty[2]
Ncediwe Mdlulwa[3]
Nondumiso Masuku[4]

AFFILIATIONS

[12&4]Human Sciences Research Council, Equitable Education and Economies, Pretoria, South Africa.
[3]Centre for Innovation in Learning and Teaching, Centre for Higher Education Development, University of Cape Town, Cape Town, South Africa.

CORRESPONDENCE

Email: psaal@hrsc.ac.za

REFERENCE

Saal, P. E., Chetty, K., Mdlulwa, N., & Masuku, N. (2026). Effective AI integration in postgraduate supervision practices: Policy implications for South Africa. In I. Kariyana & W. Sinkala (Eds.), *Artificial Intelligence and Postgraduate Supervision in Higher Education* (pp. 263-285). ERRCD Forum. https://doi.org/10.38140/obp4-2026-14



**Abstract:** Artificial Intelligence (AI) is revolutionising postgraduate supervision on a global scale by enhancing research efficiency, automating administrative tasks, and improving student engagement. Nevertheless, the adoption of AI within South African higher education, particularly in historically disadvantaged institutions, remains constrained, primarily due to a paucity of literature regarding AI adoption in postgraduate supervision. This chapter utilises the Unified Theory of Acceptance and Use of Technology (UTAUT) (Venkatesh et al., 2003) to investigate how AI tools are presently employed in postgraduate supervision across various contexts. It examines the constructs influencing the behavioural intentions of supervisors and students to adopt AI and identifies key facilitating conditions that enable its effective integration. A sectoral review was conducted employing document and thematic analysis to synthesise findings. Relevant peer-reviewed literature was sourced from traditional academic databases and AI-powered discovery tools such as SciSpace Deep Review, Elicit.com, NotebookLM, ChatGPT Deep Research, and Gemini Deep Research. Findings indicated that students predominantly utilised ChatGPT to enhance academic writing, assist with literature reviews, and receive immediate feedback, particularly when supervisors were unavailable. Conversely, supervisors employed AI to refine methodologies, data coding, and provide administrative support. Performance expectancy emerged as the most significant predictor of behavioural intention to adopt AI. However, actual adoption was contingent upon facilitating conditions such as AI literacy, peer support, institutional policies, access to infrastructure, and training opportunities. This chapter advocates for the development of comprehensive institutional frameworks to guide the ethical, pedagogical, and equitable integration of AI into postgraduate supervision.

***Keywords:*** Artificial intelligence, higher education, postgraduate supervision, South Africa.

## 1. Introduction

Artificial intelligence (AI) is rapidly transforming all spheres of education, with higher education experiencing particularly significant shifts in postgraduate supervision practices. Thong et al. (2025) noted that the integration of AI has enabled students and supervisors to engage more flexibly through virtual teaching and learning platforms, effectively reshaping traditional supervision models. In response to these developments, several studies have explored the role of AI in postgraduate education. For example, Asongo et al. (2024) and Oubibi et al. (2025)

examined how students utilise AI tools to enhance writing, conceptual understanding, and research productivity. Others, such as Chauke et al. (2024), investigated student perceptions of AI, while Caillaud and Skec (2024) focused on its ethical and methodological implications within doctoral supervision. A noteworthy contribution by Segooa et al. (2025) evaluated the use of generative AI tools to enhance the teaching of scientific research methods. Furthermore, Caillaud and Skec (2024) raised concerns about AI's potential to disrupt supervisory relationships and argued that it may act as either a valuable support mechanism or an intrusive influence. More recently, Thong et al. (2025) conducted a systematic review examining how generative AI facilitated doctoral co-supervision and found that, although AI supports personalised learning, empirical studies on AI adoption in postgraduate supervision are limited.

Other researchers have also focused on AI acceptance in postgraduate supervision, employing technology acceptance models. For instance, Sergeeva et al. (2025) conducted a quantitative study in Russia using the Unified Theory of Acceptance and Use of Technology (UTAUT) to assess generative AI adoption among university students. They identified habit and performance expectancy as key drivers of technology use. Similarly, Strzelecki (2024) employed a survey-based approach with university students in Poland to explore ChatGPT usage in higher education. This study extended the UTAUT model and found that performance expectancy, hedonic motivation, and habit significantly influenced behavioural intention. The findings highlighted generational and cultural factors that influence AI acceptance across student populations.

Despite the expanding literature on AI adoption in postgraduate supervision, several limitations remain. Much of the literature focuses narrowly on academic writing, ethics, or pedagogy and offers limited insight into how AI is integrated into supervision across disciplines and institutional contexts. While student-led adoption is well-documented (Asongo et al., 2024; Chauke et al., 2024), there is little evidence on how supervisors utilise AI or how it influences their pedagogical practices (Nikolic et al., 2024). Although some studies have applied technology acceptance models such as UTAUT (Sergeeva et al., 2025; Strzelecki, 2024), they typically focus on students and overlook supervisory dynamics. This gap is especially evident in low- and middle-income contexts such as South Africa, where infrastructural constraints, limited supervisor training, and the absence of formal policies hinder meaningful adoption (Mbangeleli & Funda, 2024).

### 1.1 Background literature review

Generative AI presents compelling arguments for its integration into postgraduate supervision, primarily due to its capacity to process vast amounts of data and provide valuable insights that would otherwise be time-consuming to attain. Nonetheless, this integration has been characterised by fragmentation, with divergent ethical perspectives and limited institutional support within the academic community (Nikolic et al., 2024). Attitudes towards AI adoption appear to be divided between the fields of science, technology, engineering, and mathematics

(STEM) and those of the humanities, with STEM practitioners embracing AI for its efficiencies in tasks such as coding and statistical analysis.

Conversely, scholars in the humanities exhibit greater scepticism, contending that although these tools possess fluent writing capabilities, they lack adequate contextual sensitivity and are susceptible to various forms of bias (Cowling et al., 2023). Hicham et al. (2025) indicate that a notable divergence exists between AI adoption in the Global North compared to the Global South, whereby developing nations such as Morocco view the social influences of AI as a more significant predictor of AI integration. This trend is particularly evident in African countries that prioritise efficiency gains due to limited resources. In contrast, the Global North appears to focus on the ethical implications and policy frameworks surrounding AI rather than its performance.

The tension between efficiency and trust significantly shapes the experiences of both supervisors and students concerning AI adoption. When interpreting AI integration through the UTAUT framework, one can discern a complex interplay among factors such as effort expectancy, social influence, and facilitating conditions. In terms of effort expectancy, student adoption is frequently motivated by the perceived benefits of speed and language assistance. Supervisors recognise the advantages of AI in alleviating their workload, yet remain acutely aware of the potential for AI hallucinations and errors (Arbulú Ballesteros et al., 2024). This awareness has heightened the necessity for supervising AI usage, resulting in increased time commitments that exacerbate their burdens and contribute to negative perceptions of AI (Mohsin et al., 2024).

With respect to social influence, supervisors and students face differing pressures regarding AI adoption. Institutional norms, policies, and endorsements significantly shape supervisors' perspectives on AI integration, while the adoption trends among peers generate a sense of fear of missing out for students (Tao et al., 2024). Simultaneously, students' utilisation of AI is tempered by apprehensions concerning plagiarism and academic misconduct. Fear also plays a role in influencing supervisors' adoption decisions, particularly regarding concerns about AI potentially replacing human roles and threatening their epistemic authority (Khlaif et al., 2024).

Regarding the facilitating conditions for AI adoption within the South African context, digital inequality has been identified as a significant driver of AI adoption. Those with access to digital infrastructure and the financial means to afford the data required for the application of AI tools are better positioned to implement such technologies (Hicham et al., 2025). Furthermore, the policies that regulate AI adoption provide essential frameworks for institutions to navigate the associated tensions, ensuring that adoption is executed ethically and accompanied by appropriate training for both staff and students (Nikolic et al., 2024).

The UTAUT framework also promotes an analysis of AI adoption from both individual and organisational perspectives, while recognising the technical factors that influence adoption. Among supervisors, Fang et al. (2025) contend that personal attitudes towards AI can significantly drive adoption. By fostering an environment of openness and demonstrating the benefits of innovation and self-efficacy, supervisors may come to see AI as an instrument for enhancement rather than a threat. Such demonstrations are feasible when student experimentation leads to innovations that stimulate curiosity, learning, and research (Tian et al., 2024). From an organisational perspective, as posited by Nikolic et al. (2024), AI policies serve as crucial frameworks that can cultivate a culture of innovation and facilitate successful adoption. Hicham et al. (2025) found that these policies were the most pivotal factor influencing adoption in Moroccan universities. Importantly, these policies must be complemented by funding and training to ensure that institutions are adequately equipped to operationalise AI tools effectively (Baharin et al., 2025).

AI adoption is also influenced by technical factors pertaining to its accessibility and reliability. Premium generative AI tools frequently necessitate subscriptions, contributing to unequal diffusion among supervisors and students. Therefore, policies are required to ensure that the benefits of these tools are distributed equitably (Mosae & Kaushal, 2025). Additionally, perceived security is a critical determinant of AI acceptance; thus, policies must clearly demonstrate how data can be protected, ensuring that neither user nor research participant information is compromised (Ratta et al., 2025). In this context, when these factors are aligned, Mohsin et al. (2024) assert that AI adoption has the potential to enhance supervisory efficiency and personalisation, consequently improving student outcomes and academic oversight.

### 1.2 Problem statement

This chapter seeks to fill the gaps in the literature by investigating how AI is currently being used in postgraduate supervision across disciplines and countries, guided by the UTAUT framework (Venkatesh et al., 2003). It explores the factors influencing the behavioural intentions of supervisors and students to adopt AI, identifies the individual, organisational, and technical conditions that facilitate actual use, and analyses the policy implications for integrating AI into supervision within the South African higher education system. This qualitative, document-based study draws on academic literature and employs document and thematic analysis to synthesise findings.

While the literature review, through a cursory examination of available studies, presents several factors influencing AI adoption among supervisors and students, these factors remain disparate and fragmented. A synthesis of these factors is needed to explain their interaction within a coherent theoretical framework. These views are sourced from across the Global South and North and have implications for the South African context, but they need to be reviewed in light of South African inequality and resource constraints. Thus, there is a need for a comprehensive, theoretically grounded examination of AI adoption in postgraduate supervision that consolidates existing evidence, identifies patterns across disciplines and geographies, and derives actionable policy recommendations. The research questions addressed are:

- How is AI being used in postgraduate supervision across different disciplines and countries?
- Which factors influence the behavioural intention of supervisors and students to use AI in supervision practices?
- Which individual, organisational, and technical factors mediate the adoption of AI in postgraduate supervision?
- What policy implications can be drawn to guide effective AI integration in South African higher education?

The next section introduces the theoretical framework used in the study, followed by the methods employed. Thereafter, the findings are presented and discussed in relation to the research questions. Lastly, this chapter concludes with practical recommendations for policy and practice.

## 2. Theoretical Framework

This study is grounded in the UTAUT, one of the most widely applied models for understanding technology adoption in educational settings (Venkatesh et al., 2003). The UTAUT model focuses on four key constructs, namely, performance expectancy, effort expectancy, social influence, and facilitating conditions. Performance expectancy is defined as "the degree to which an individual believes that using the system will help him or her to attain gains in job performance" (Venkatesh et al., 2003, p. 447). In this chapter, it refers to the belief held by students and supervisors that AI enhances research and learning. Effort expectancy is described as "the degree of ease associated with using the system" (Venkatesh et al., 2003, p. 450), and we define it as AI tools that require little effort to use. Social influence is "the degree to which an individual perceives that important others believe he or she should use the new system" (Venkatesh et al., 2003, p. 451). We refer to this as the extent to which supervisors and students believe that significant individuals in their lives think they should use AI. These constructs influence users' behavioural intention to use technology, which may ultimately determine their actual usage. Furthermore, this chapter focuses on facilitating conditions, defined as "the degree to which an individual believes that an organisational and technical infrastructure exists to support the use of the system" (Venkatesh et al., 2003, p. 450). In this chapter, we describe facilitating conditions as individual, organisational, and technical factors that mediate the effective adoption of AI in postgraduate supervision practices.

## 3. Methodology

We employed a sectoral literature review to investigate the utilisation of AI in postgraduate supervision, the constructs influencing users' behavioural intent to employ AI, and the factors facilitating its use across various disciplines and countries. We conducted an extensive document analysis guided by a structured deep-search strategy (Bowen, 2009). This involved utilising traditional academic databases, such as Google Scholar, in conjunction with AI-powered semantic search tools, including SciSpace Deep Review, Elicit.com, NotebookLM, ChatGPT Deep Research, and Gemini Deep Research. These tools enabled us to identify and retrieve academic literature pertinent to the scope of the study. Following the document analysis, we employed thematic analysis to systematically categorise the findings (Braun & Clarke, 2006).

Initially, we searched for literature using key phrases derived from the research questions. The traditional literature search was supplemented by a selection of studies sourced from the aforementioned AI-powered semantic search tools. Specifically, we sought case studies and empirical accounts of AI adoption in higher education settings, with particular attention to studies that capture the experiences, perceptions, and practices of postgraduate supervisors. We continued the iterative search process until we reached theoretical saturation, the point at which no new themes or substantive insights emerged from additional sources. However, while saturation was achieved for broader themes related to AI in higher education, the evidence base for supervision practices remains limited.

The findings were first organised according to the research questions. We grouped the findings from research question one based on emerging themes. Subsequently, we categorised research question two, focusing on the UTAUT constructs according to predetermined themes (performance expectancy, effort expectancy, social influence). Research question three, which centred on facilitating conditions, was organised into predetermined themes (individual, organisational, and technical levels).

To ensure the trustworthiness of the study, we applied strategies aligned with qualitative research principles, specifically focusing on credibility and confirmability (O'Leary, 2017). Credibility was maintained through the systematic use of multiple high-quality data sources from peer-reviewed journals, academic databases, and AI-assisted platforms. We carefully assessed each source obtained from SciSpace Deep Review, Elicit.com, NotebookLM, ChatGPT Deep Research, and Gemini Deep Research to ensure its relevance and accuracy. Confirmability was strengthened through collaborative review and cross-validation of findings among the research team, ensuring that interpretations were grounded in evidence rather than personal bias.

## 4. Findings

In the subsequent sections, we address the first three research questions. We discuss the last research question in the discussion section, focusing on policy implications for South Africa to avoid repetition.

### 4.1 AI Uses in postgraduate supervision across disciplines and countries

In this section, we present the actual applications of AI in postgraduate supervision practices across various national and disciplinary contexts (see Tables 1a-e). Although an increasing number of publications underscore the potential of AI in higher education, we identified relatively few empirical studies that concentrate specifically on its implementation within postgraduate supervision. This limited evidence base may be attributed to the recent emergence of generative AI tools in academic environments.

**Table 1a:** *Applications of AI by postgraduate students in research and writing*

| Country | Discipline/field | AI Tool(s) | AI Use | Citation |
|---|---|---|---|---|
| Jordan | English First Language (discipline unspecified) | ChatGPT | Thesis writing | Amer et al. (2025) |
| Nigeria | Discipline not specified | ChatGPT, Quilbot, ChatPDF, Consensus, Scite, Bit AI, Litmap, Jenni, Paperpal, Research Rabbit, Wordvice AI, Typeset.io | Enhance research | Asongo et al. (2024) |
| China | Education | ChatGPT, Grammarly; QuillBot | Academic writing | Oubibi et al. (2025) |
| South Africa | Human, Social Science and Education | ChatGPT | Refine research topics before submission to supervisors; assist with paraphrasing enhance academic writing, especially for non-native English speakers; formulate research ideas and problem statements; search for relevant literature and write literature reviews. | Chauke et al. (2024) |
| China | Medical education/ Health | ChatGPT and other tools | As conversational tools for studying — used by students to quickly obtain medical information and knowledge. | Tao et al. (2024). |
| Peru | All fields | ChatGPT | AI as a learning support tool. | Arbulú Ballesteros et al. (2024). |

**Table 1b:** *Applications of AI in postgraduate supervision and mentorship*

| Country | Discipline/field | AI Tool(s) | AI Use | Citation |
|---|---|---|---|---|
| Australia | Education/Research Supervision | ChatGPT | AI is used to support psychological need fulfilment, autonomy, competence, and relatedness among research students through preliminary formative feedback, literature review assistance, idea generation, proofreading, and editing. | Cowling et al. (2023) |
| France; Sweden; Italy; Slovenia; and Croatia | Engineering Design | ChatGPT, Scholar AI | PhD Supervisors: Finding relevant papers and supporting literature review; Summarising and synthesising; Analysing data and coding; Improving research methodology; Writing task such | Caillaud and Skec (2024) |

| Country | Discipline/field | AI Tool(s) | AI Use | Citation |
|---|---|---|---|---|
| | | | as administrative and routine tasks | |
| United Kingdom | Not specified | ChatGPT (GPT4) | Fill the gaps left by the supervisor; Provide an interactive, private space; Test and refine complex research questions. Facilitate a deeper understanding of the research area. Conceptualise research focus; Generate new ideas; Literature review. advance understanding of methodological designs; Reflect on ethical considerations. Grammatical and structural checks | Harding and Boyd (2024) |
| Uganda | Computer Sciences and Electronics, Natural Sciences, Social Sciences, Education, Arts and Humanities | All generative AI tools | Students use AI to generate research content, which supervisors struggle to detect due to their limited knowledge of AI tools. | Rajab et al. (2025) |

***Table 1c:*** *AI Integration in Postgraduate Teaching and Curriculum Design*

| Country | Discipline/field | AI Tool(s) | AI Use | Citation |
|---|---|---|---|---|
| Malaysia | General | ChatGPT and other generative AI tools | AI is used to enhance postgraduate teaching by providing personalised learning experiences, fostering active engagement, and supporting educators in identifying suitable learning theories and strategies | Omar et al. (2024) |
| Not mentioned | Pathology and microbiology | Generative AI tools in general | AI is used to provide personalised learning experiences, simulate complex scenarios, automate assessments, and enhance teaching performance in postgraduate supervision | Roy et al. (2024) |
| Global | Medical education | Generative AI tools in general | AI is used in postgraduate medical education to provide personalised learning experiences, real-time feedback, AI mentorship, and assistance in medical research and evidence-based practices. | Pashkovskyy et al. (2023) |
| Saudi Arabia | Education | Generative AI tools in general | AI is used in scientific writing by postgraduate students to enhance academic integrity, detect plagiarism, facilitate automated reviews , improve writing skills , and aid in formatting scientific papers. | Hegazy et al. (2024) |

***Table 1d:*** *Global perspectives and attitudes toward AI in higher education*

| Country | Discipline/field | AI Tool(s) | AI Use | Citation |
|---|---|---|---|---|
| Global | All fields | Generative AI tools in general | Studied in terms of teachers' attitudes, perceptions, intentions and behaviours toward using AI/GenAI | Nikolic et al. (2024) |
| Morocco | Social sciences, exact sciences and medical disciplines | Large language models | Examined as a technology whose adoption intention is influenced by factors like social influence and facilitating conditions. | Hicham et al. (2025) |
| Malaysia | All fields | Generative AI tools in general | Examine attitudes and behavioural intentions to adopt AI in their education based on UTAUT factors. | Mohsin et al. (2024). |
| Middle East | All fields | Generative AI tools in general | Staff integrating AI into assessment practices to design assignments and manage workloads. | Khlaif et al. (2024). |
| China | Library and Information Science | Generative AI tools in general | Examines librarians' intention to adopt AI in library routines and services | Fang et al. (2025) |
| Global | All fields | ChatGPT, Grammarly | As learning support and writing/editing aid tools that simplify complex content, improve writing and grammar, support personalised learning, and speed up research. | Mosae & Kaushal (2025) |

***Table 1e:*** *General and cross-disciplinary applications of AI in postgraduate education and research*

| Country | Discipline/field | AI Tool(s) | AI Use | Citation |
|---|---|---|---|---|
| Australia | Postgraduate Research (General) | ChatGPT | Personalised tutoring; Language editing and proofreading; Brainstorming, Coding and interpretation; Literature synthesis; Mock interaction and rehearsal | Dai et al. (2023) |
| Global | All fields | Generative AI tools in general | AI is used to personalise teaching, provide formative feedback, identify at-risk students, accelerate research discovery, streamline administrative processes, and optimise resource utilisation. | Tarisayi (2024) |
| China | Language Education/Linguistics | EvaluMate | AI is used to support peer review through an AI-powered system called EvaluMate, which includes a chatbot named Eva that evaluates and provides feedback on student reviewers' comments to improve the quality of peer feedback. | Guo et al. (2025) |
| China | All fields | AI Chatbots | Used as conversational technology for study and research support. | Tian et al. (2024) |
| India | Medical sciences | AI-driven Clinical Decision Support Systems | As decision-support tools for clinicians | Ratta et al. (2025) |

### *4.1.1 Academic writing and language support*

A significant area of AI use involves support for academic writing and language development. According to Oubibi et al. (2025), the integration of AI in higher education institutions is recognised as a transformative tool that enhances academic writing for postgraduate students, effectively addressing their writing challenges. In Jordan, Amer et al. (2025) noted that students used ChatGPT to enhance thesis writing, particularly to expand vocabulary, clarify meaning, and generate coherent text. Similarly, in South Africa, students relied on ChatGPT to paraphrase, restructure sentences, and improve the grammatical quality of their work before submitting it to supervisors (Chauke et al., 2024). In Nigeria, Asongo et al. (2024) found that students used ChatGPT and Quillbot to improve sentence clarity and coherence, while in China, students reported using ChatGPT, Grammarly, and Quillbot to check grammar, refine sentence structure, and edit formal writing (Oubibi et al., 2025). Additionally, in a study focusing on pathology and microbiology, generative AI was noted for its potential to grade written assignments, reduce faculty workload, and provide students with formative feedback on their scientific writing (Roy et al., 2024).

### *4.1.2 Research development and conceptualisation*

AI was also used to support the early stages of research development. Large Language Models (LLMs) like ChatGPT are uniquely suited to assist students in the beginning stages of their candidature by helping them articulate key concepts through generated text, allowing them to focus more on the concepts themselves than on the written expression of ideas (Cowling et al., 2023). In South Africa, students used ChatGPT to clarify research objectives, develop problem statements, and refine their study focus (Chauke et al., 2024). In Australia, Dai et al. (2023) reported that students used the tool for brainstorming, conceptual planning, and refining research topics.

The United Kingdom presented similar findings, where students engaged with ChatGPT-4 to test and reformulate complex research questions and enhance the clarity of their theoretical framing (Harding & Boyd, 2024). Similarly, the study conducted by Dai et al. (2023) examined the use of ChatGPT and its impact on postgraduate research supervision. Insights gathered from Australian postgraduate research students indicated that ChatGPT not only enhanced research and academic performance but also fostered critical thinking skills. Students reported that the AI generative tool allowed them to work more independently and encouraged self-directed learning in their research endeavours.

### *4.1.3 Literature review and source discovery*

In addition to writing and conceptualisation, both students and certain supervisors have utilised artificial intelligence (AI) to facilitate literature review tasks. AI algorithms can support postgraduate students in researching and analysing extensive volumes of academic literature by employing Natural Language Processing (NLP) techniques to extract pertinent information, summarise research articles, and provide evidence-based recommendations (Pashkovskyy et al., 2023). In Nigeria, students have employed tools such as Litmaps, Scite, and Elicit to identify relevant literature, explore citation networks, and construct structured reviews (Asongo et al., 2024). In South Africa, students have utilised ChatGPT to define key search terms and summarise foundational texts (Chauke et al., 2024). Furthermore, supervisors in France, Sweden, Italy, Slovenia, and Croatia have also employed AI tools such as Scholar AI and ChatGPT to assist doctoral candidates in locating pertinent publications and synthesising academic arguments (Caillaud & Skec, 2024).

#### *4.1.4 Methodological and ethical reflection*

Beyond topic development and access to literature, AI has supported students in engaging with research methodology and ethical design. However, a core limitation is that AI language models, such as ChatGPT, lack a specific ethical framework and operate based on the data and algorithms on which they were trained. This can result in biased or flawed advice, making critical review by the student essential (Hegazy et al., 2024). In Australia, students have consulted ChatGPT for assistance with interpreting methodological texts, understanding research designs, and reflecting on ethical concerns related to data collection (Dai et al., 2023). In European contexts, particularly in France and Italy, supervisors have employed AI to assist students in refining methodological frameworks and comparing the strengths of various design approaches (Caillaud & Skec, 2024).

#### *4.1.5 Personalised tutoring and independent learning*

AI also provided a form of personalised academic support, particularly in environments where students had limited access to supervisors. Generative AI can significantly enhance learning experiences by offering interactive and personalised educational tools, such as generating realistic images or simulating complex processes, and by creating virtual tutors that provide personalised feedback and guidance. The integration of AI into postgraduate medical education facilitates a more personalised and adaptive learning experience by analysing individual learning patterns and recommending tailored study plans (Pashkovskyy et al., 2023). In Australia, Dai et al. (2023) found that students used ChatGPT as a private tutor to review language, provide feedback on drafts, help interpret data, and rehearse academic presentations. Similarly, in China, students utilised ChatGPT for idea generation, editing, and practising responses to critical questions (Oubibi et al., 2025).

#### *4.1.6 Filling supervisory gaps and supporting reflection*

In settings where supervision was inconsistent or limited, students used AI to bridge the gap. For research students, ChatGPT can provide preliminary formative feedback, checking for errors and critiquing early ideas. This enables students to have higher-impact conversations with their human supervisors (Cowling et al., 2023; Hegazy et al., 2024). In the United Kingdom, Harding and Boyd (2024) found that ChatGPT-4 served as a reflective partner, helping students rehearse arguments, clarify complex ideas, and receive immediate feedback. Students described the tool as a non-judgmental space for intellectual experimentation, which allowed them to progress even in the absence of timely supervisor input. Similarly, in South Africa, students reported that AI supported their progress between supervision sessions, particularly when institutional resources were stretched (Chauke et al., 2024).

### *4.1.7 Supervisor use of AI for academic and administrative tasks*

Although supervisors used AI less frequently than students, some applications were identified. AI is being used to streamline administrative processes in higher education through chatbots, and it has the potential to automate administrative teaching tasks, identify areas needing classroom reinforcement, and intelligently utilise data for student support (Rajab et al., 2025). Furthermore, AI can support educators' professional development by offering teaching evaluation models and suggestions to enhance their instructional practices (Omar et al., 2024). In France, Sweden, Italy, Slovenia, and Croatia, Caillaud and Skec (2024) found that supervisors employed AI tools to assist in identifying relevant articles, coding qualitative data, and refining research methodology sections. Additionally, supervisors reported using AI to support administrative tasks such as writing feedback summaries and managing supervision documents.

## 4.2 Factors influencing the behavioural intention to Use AI in postgraduate supervision

In this section, we discuss how performance expectancy, effort expectancy, and social influence, as stated in the UTUAT (Venkatesh et al., 2003), influenced the supervisors' and students' behavioural intention to use AI in postgraduate supervision practices.

### *4.2.1 Performance expectancy*

This construct emerged as a strong predictor of the intention to adopt AI. For instance, Acosta-Enriquez et al. (2024) found in their systematic review of AI acceptance in universities that performance expectancy was a significant predictor of students' and faculty's intention to use AI. The potential for AI to enhance postgraduate teaching is realised by promoting active engagement and meaningful learning interactions, thereby improving learning outcomes, productivity, and student engagement through personalised education, feedback, and assistance (Omar et al., 2024). Performance expectancy was evident in how students perceived AI tools such as ChatGPT and Grammarly to enhance their academic output, particularly in supporting research writing, conceptual development, and methodological clarity. For instance, students in education and the humanities viewed these tools as useful for managing their academic workload and producing higher-quality work (Chauke et al., 2024; Dai et al., 2023). Supervisors also believed that ChatGPT could enhance thesis writing, vocabulary usage, and flow (Amer et al., 2025).

### *4.2.2 Effort expectancy*

We found limited evidence regarding effort expectancy, which we refer to as AI not being difficult to use. Although many instructors are unaware of AI's full potential and underlying principles, the utilisation of AI in education has led to the complete integration of teaching and learning, suggesting it is readily accessible. However, its successful adoption is influenced by users' level of trust in the tools (Hegazy et al., 2024; Roy et al., 2024). We report on how students perceived the use of AI, given that limited supervisors employed AI tools in their practices. Acosta-Enriquez et al. (2024) found that perceived ease of use emerged as a relevant factor in students' AI adoption decisions. Furthermore, Patterson et al. (2024) reported that students showed high agreement that AI tools were easy to use. Amer et al. (2025) also recommended training and upskilling both supervisors and students in AI functionalities, indicating that some effort is needed to adopt AI in practice. In contrast, Aljarboa et al. (2025) found that effort efficiency does not influence students' behavioural intention to adopt AI across disciplines, including science, engineering, business, and the humanities, in Saudi Arabia.

### *4.2.3 Social influence*

Cowling et al. (2023) highlighted the importance of leadership in fostering AI integration within supervision. Their study found that students appreciated the use of ChatGPT to support their research productivity and academic empowerment. At the departmental level, leadership practices created opportunities for formative feedback, making AI tools more accessible and useful to both undergraduate and postgraduate students. However, the study also noted that leadership influence was sometimes limited to reinforcing existing norms rather than cultivating forward-looking practices. Nonetheless, the presence of supportive academic leadership contributed meaningfully to the social dynamics that encouraged AI adoption in supervision contexts (Tarisayi, 2024). Informal peer learning networks in countries such as Nigeria and South Africa played a pivotal role in normalising the use of AI, while departmental cultures in parts of Europe fostered collective engagement with emerging technologies (Asongo et al., 2024; Caillaud & Skec, 2024; Chauke et al., 2024). Furthermore, Supianto et al. (2024) found, using structural equation modelling, a positive correlation between social influence and Indonesian students' behavioural intention to use ChatGPT. This indicates that when Indonesian students feel encouraged or supported by their peers and significant others to use ChatGPT, their intention to adopt this technology increases.

## 4.3 Enabling conditions for AI integration in postgraduate supervision

In the previous sections, we discussed the constructs that influenced behavioural intention. However, even when students and supervisors intend to use AI, they will only employ AI in postgraduate supervision practices if certain facilitating conditions are in place. Table 2 summarises the factors that support the use of AI across contexts at the individual, organisational, and technical levels.

**Table 2.** *Factors facilitating the adoption of AI across countries and disciplines*

| Level | Facilitating condition | Citation |
|---|---|---|
| Individual | AI literacy<br>Peer support | Asongo et al. (2024); Gandhi et al. (2024);<br>Asongo et al. (2024); Chauke et al. (2024)<br>Caillaud and Skec (2024) |
| Organisational | Policies and guidelines | Amer et al. (2025); Caillaud and Skec (2024); Chauke et al. (2024); Dai et al. (2023) |
| | Training and workshops | Acosta-Enriquez et al. (2024); Amer et al. (2025) |
| Technical | Reliable access to AI tools and infrastructure | Asongo et al. (2024); Chauke et al. 2024; Oubibi et al. 2025; Habibi et al. 2023 |

### *4.3.1 Individual-level facilitating factors*

A key facilitating condition at the individual level is the capability and readiness of supervisors and students to use AI. A supervisor or student might not use AI in practice if they lack the knowledge or skills to do so. For instance, Gandhi et al. (2024) conducted a mixed-methods study to compare AI knowledge among medical students and found that postgraduates with greater AI knowledge engaged in more AI-related practices, while those with lower AI knowledge used AI less. The authors emphasised the need for targeted AI education to prepare students. A study by Asongo et al. (2024) acknowledged the limited research on postgraduate students' awareness and utilisation of AI technologies for research purposes. Asongo et al. (2024) argue that integrating AI technologies into higher education can provide researchers with deeper insights into learning processes and effective teaching methods. Furthermore, supervisors interviewed by Caillaud and Skec (2024) expressed a need to learn how to use AI ethically. This concern may stem from the fear that students might soon outpace their supervisors in AI proficiency (Caillaud & Skec, 2024).

In Nigeria, Asongo et al. (2024) found that peer networks served as an informal support system for learning about AI tools. Students relied on each other to share strategies and recommend tools for academic writing and research. Similarly, Chauke et al. (2024) reported that students in South Africa benefited from informal peer knowledge, especially in the absence of institutional training. These findings align with Caillaud and Skec (2024), who observed that peer and departmental culture influenced the legitimacy of AI use in European doctoral programmes. As some peers embraced AI, others felt more comfortable doing the same.

### *4.3.2 Organisational factors*

Several studies emphasised the critical need for higher learning institutions to develop clear policies, guidelines, and regulations on the acceptable and ethical use of AI tools like ChatGPT in postgraduate research and thesis writing (Amer et al., 2025; Caillaud & Skec, 2024; Chauke et al., 2024; Dai et al., 2023). For example, in the study by Caillaud and Skec (2024), supervisors and students from Europe (France, Sweden, Italy, Slovenia, and Croatia) highlighted the need for institutional regulations and explicit guidelines to ensure the proper and ethically acceptable use of generative AI. We also found that in South Africa, the University of Pretoria's Faculty of Health Sciences is guiding responsible AI use for postgraduate research (University of Pretoria, Faculty of Health Sciences, 2024). The institution has also provided a lecturer's guide, which offers practical strategies for the effective and ethical integration of generative AI into teaching (University of Pretoria, Faculty of Health Sciences, 2024). However, no provision is made for postgraduate supervision across disciplines. Researchers argue that contextualised policies and guidelines would set boundaries for AI use and address concerns related to academic integrity and responsible integration (Amer et al., 2025; Caillaud & Skec, 2024; Chauke et al., 2024).

Beyond individual self-learning, institutional training programmes can greatly support adoption. Universities that offer workshops on AI tools for research, provide documentation or helpdesks for AI software, or include AI modules in researcher development courses effectively equip their members to use AI. The literature frequently calls for such organisational initiatives. This construct also emerged as a strong predictor of the intention to adopt AI. For instance, Acosta-Enriquez et al. (2024), in their systematic review of AI acceptance in universities, suggest that training should be emphasised as an organisational strategy to leverage AI use. Additionally, Amer et al. (2025) highlighted that training is needed to use ChatGPT optimally.

### *4.3.3 Technical factors*

Access to digital infrastructure, including personal devices, connectivity, and free online platforms, was a foundational enabling condition. In South Africa, Chauke et al. (2024) found that students used mobile phones and mobile data to access ChatGPT in the absence of institutional infrastructure. Similarly, in Nigeria, Asongo et al. (2024) reported that students utilised free AI tools, such as ChatGPT, Quillbot, and Elicit, without institutional licences. In China, Oubibi et al. (2025) observed a similar use of ChatGPT and Grammarly, both accessed externally from the university. The availability and affordability of these tools enabled students to integrate AI into their academic work, regardless of formal provision. In Indonesia, Habibi et al. (2023) identified facilitating conditions, including support for tools and infrastructure, as the most significant predictors of students' intention to adopt AI in learning.

## 8. Discussion of Findings

This study examined the factors that influence students' and supervisors' behavioural intentions to adopt AI, the facilitating conditions that enable actual use, and the policy implications for effective AI integration, particularly within South African higher education. The findings are discussed in relation to the four research questions that framed this study.

AI tools are increasingly being integrated to streamline various research activities. However, usage patterns differ between students and supervisors. Students leverage AI tools to initiate their research, particularly for brainstorming topics and structuring initial drafts. In contrast, supervisors adopt a more cautious approach, using AI selectively to support tasks such as refining research methodologies, conducting literature searches, coding qualitative data, and managing administrative processes (Caillaud & Skec, 2024). Tools such as ChatGPT, Grammarly, and Quillbot are widely utilised to enhance grammatical accuracy, optimise clarity, and overcome linguistic barriers, particularly among students for whom English is a second or additional language (Chauke et al., 2024; Asongo et al., 2024). This observation aligns with the growing literature emphasising that AI tools are democratising academic writing by providing language support for English second-language learners (Wang & Wang, 2025).

The findings indicate that multiple factors shape the decision to adopt AI tools. The perceived usefulness of AI tools in expediting the research process emerges as a key driver of integration in academic settings. This supports the UTAUT model, which links performance benefits to technology uptake. Wang and Wang (2025) and Acosta-Enriquez et al. (2024) confirm that students value AI for speeding up research tasks and providing timely feedback, addressing gaps that traditional supervision cannot fill (Dai et al., 2023; Chauke et al., 2024). Performance expectancy emerges as the strongest predictor of AI use in higher education (Patterson et al., 2024).

Regarding the effort required to utilise AI tools, students generally find these tools user-friendly. Caillaud and Skec (2024) observe that while supervisors have some experience with AI tools, they often lack practical proficiency to apply them across various research activities. Effort expectancy plays a less significant role in academic writing contexts (Patterson et al., 2024). Despite widespread access, structured support remains essential to ensure the responsible and effective use of AI in academic writing (Ngoc et al., 2025).

Social influence emerges as another key driver of AI adoption in academic writing. In research environments where the responsible use of AI is endorsed by peers, students are more inclined to adopt AI tools (Funda & Piderit, 2024; Caillaud & Skec, 2024). Al-Bukhrani et al. (2025) argue that when influential figures, such as peers, supervisors, and the broader research setting, promote AI integration, perceived barriers are reduced, thereby encouraging student engagement. Supianto et al. (2024) also report a positive relationship between social influence and behavioural intention, reinforcing the crucial role of social factors in shaping students' willingness to use AI tools in their research. Cowling et al. (2023) emphasise that universities and research supervisors build students' competence in using AI tools by promoting digital literacy and guiding effective use in academic writing.

Individual-level facilitating factors can strongly predict AI integration in academic settings, particularly when supervisors and students are equipped with digital literacy and targeted AI training to use the tools effectively. However, any perceived difficulty in using these tools can negatively influence adoption decisions (Gandhi et al., 2024; Asongo et al., 2024; Caillaud &

Skec, 2024). In South Africa and Nigeria, informal networks have often compensated for the lack of institutional training, driving students' willingness to adopt AI tools (Asongo et al., 2024; Chauke et al., 2024). At the organisational level, institutions can promote AI adoption by implementing clear policies and guidelines on responsible use, alongside offering training and support to students (Acosta-Enriquez et al., 2024; Amer et al., 2025). Finally, technical factors, such as access to digital infrastructure and free online AI tools, are foundational. Even in the absence of institutional provision, students across Nigeria, South Africa, China, and Indonesia have used AI for research, highlighting the critical role of accessible technology and infrastructure in driving adoption (Asongo et al., 2024; Chauke et al., 2024; Oubibi et al., 2025; Habibi et al., 2023).

In Northern Nigeria, particularly in Benue State, many postgraduate students struggle with academic writing, information sourcing, and producing original work (Asongo et al., 2024). Consequently, this study proposes several policy recommendations for effective AI integration in higher education. It is suggested that faculty and departmental heads prioritise organising seminars and workshops to educate postgraduate students on the effective use of various AI tools to enhance their research skills (Asongo et al., 2024). Effective AI integration in higher education can be further improved through the development of students' technical and digital literacy skills, such as training students to use AI tools effectively and ethically in their academic endeavours (Oubibi et al., 2025). Thus, developing these skills helps postgraduate students utilise digital tools effectively, promoting academic engagement and improved writing (Oubibi et al., 2025).

Similarly, in South Africa, AI integration into postgraduate supervision remains largely informal and student-driven. One key implication is the necessity of establishing clear institutional guidelines on the ethical and pedagogical use of AI in research, as informal approaches risk deepening inequalities in access and usage (Chauke et al., 2024). Therefore, universities should develop comprehensive AI integration frameworks that include training modules, codes of conduct, and assessment guidelines to support both students and academic staff. Additionally, the digital divide across South African higher education institutions, particularly between historically advantaged and disadvantaged universities, calls for differentiated policy responses. Targeted investment in digital infrastructure, subsidised access to AI tools, and the inclusion of AI literacy within postgraduate orientation and research methods curricula are essential. Faculty development should also be prioritised.

Furthermore, institutional support must be extended to create a culture that promotes responsible innovation. Policies should encourage interdisciplinary collaboration and the formation of AI learning communities within faculties. These communities can serve as peer support platforms that normalise and disseminate best practices in AI usage. Finally, given the benefits of AI in addressing language barriers and promoting autonomous learning (Chauke et al., 2024; Oubibi et al., 2025), it is essential that South African higher education policies integrate

AI tools into broader academic support systems. This is particularly important for students from multilingual and historically under-resourced institutions, where language and access challenges continue to affect research participation and success.

## 9. Conclusions and Recommendations

The purpose of this study was to investigate the adoption of AI in postgraduate supervision across various disciplines and countries, utilising the UTAUT (Venkatesh et al., 2003) as a guiding framework. Furthermore, we examined how the constructs of UTAUT, particularly performance expectancy, effort expectancy, and social influence, influenced the behavioural intentions of postgraduate students and supervisors regarding the adoption of AI tools. Additionally, we identified the key facilitating conditions necessary for transitioning from intention to actual use and assessed the policy implications for effective AI integration into postgraduate supervision practices within South African higher education.

The findings indicated that performance expectancy was a strong predictor of the intentions of both supervisors and students to adopt AI in postgraduate supervision practices. Moreover, facilitating conditions such as AI literacy, peer support, clear institutional policies, training, and reliable access to infrastructure were identified as critical factors for actual adoption. Notably, it was observed that students, rather than supervisors, were at the forefront of the integration of AI into postgraduate supervision practices. Furthermore, AI was increasingly employed to address supervision gaps by providing students with immediate feedback and intellectual companionship when formal supervisory support was limited.

This study carries several social and practical implications. Socially, the findings underscore the potential for AI to democratise access to academic support, particularly for students in resource-constrained environments or those for whom English is an additional language. By bridging supervisory gaps and offering immediate feedback, AI tools can foster greater equity in postgraduate education. Practically, the study highlights the urgent necessity for South African higher education institutions to progress beyond informal, student-driven AI adoption towards the establishment of structured institutional frameworks that ensure the ethical, equitable, and pedagogically sound integration of AI into supervision practices.

Based on this, the following recommendations were made:

- Future researchers could conduct a qualitative study to investigate the limited use of AI by supervisors.
- Research could also focus on supervisors' perspectives, exploring their training needs and barriers to AI adoption.
- We recommend the development of comprehensive institutional frameworks that govern the ethical, pedagogical, and practical use of AI in postgraduate education.
- Universities should also establish clear guidelines, training programmes, and resource support systems that promote responsible AI use.

### 9.1 Limitations of the study

While this chapter provides valuable insights into the global utilisation of AI in postgraduate supervision, the constructs influencing behavioural intent, and the factors facilitating its implementation in these practices, it also highlights several limitations. Firstly, the research relied exclusively on secondary data, which, although rigorously sourced and reviewed, restricted the inclusion of practice-based experiences from supervisors and students. Secondly, the algorithms employed in generative AI tools may have influenced source selection through inherent biases or content prioritisation. Lastly, the absence of empirical data collection hindered the study's ability to triangulate findings with an alternative stream of data, which could have enhanced the contextual understanding of AI integration in supervisory practices.

## 10 Declarations

**Funding:** This research did not receive any external funding.
**Conflicts of Interest:** The authors declare no conflict of interest.
**Use of Artificial Intelligence:** The current work was created with the assistance of artificial intelligence technologies (NotebookLM, Elicit.com, Scispace Deep review, ChatGPT Deep Research and Gemini Deep Research) to assist with locating, organising the literature and identifying common themes; language and grammar clarifications; and aligning references according to APA 7th, as confirmed by the authors.

## References

Acosta-Enríquez, B. G., Ramos Farroñan, E. V., Villena Zapata, L. I., Mogollón Garcia, F. S., Rabanal-León, H. C., Morales Angaspilco, J. E., & Saldaña Bocanegra, J. C. (2024). Acceptance of artificial intelligence in university contexts: A conceptual analysis based on UTAUT2 theory. *Heliyon, 10*(19), e38315. https://doi.org/10.1016/j.heliyon.2024.e38315

Al-Bukhrani, M. A., Alrefaee, Y. M. H., & Tawfik, M. (2025). Adoption of AI writing tools among academic researchers: A theory of reasoned action approach. *PLOS ONE, 20*(1), e0313837. https://doi.org/10.1371/journal.pone.0313837

Aljarboa, S., Alabdulatif, A., & Bashir, M. (2025). The adoption of artificial intelligence for higher education. *Fusion: Practice and Applications, 20*(1), 55–67. https://doi.org/10.54216/FPA.200105

Amer, M. L. B., Rababah, L. M., & Rababah, M. A. (2025). Exploring ChatGPT as a tool for thesis writing: Perspectives of EFL supervisors in Jordanian universities. *Journal of Language Teaching and Research, 16*(1), 33–42. https://doi.org/10.17507/jltr.1601.04

Arbulú Ballesteros, M. A., Acosta Enríquez, B. G., Ramos Farroñán, E. V., García Juárez, H. D., Cruz Salinas, L. E., Blas Sánchez, J. E., Arbulú Castillo, J. C., Licapa-Redolfo, G. S., & Farfán Chilicaus, G. C. (2024). The sustainable integration of AI in higher education: Analysing ChatGPT acceptance factors through an extended UTAUT2 framework in Peruvian universities. *Sustainability, 16*(23), 10707. https://doi.org/10.3390/su162310707

Asongo, T. S., Akuse, S. S., & Aza, I. (2024). Awareness and utilisation of artificial intelligence (AI) tools for enhanced research among postgraduate students in universities in Benue State. *International Journal of Innovative Science and Research Technology, 9*(9), 1712–1720. https://doi.org/10.38124/ijisrt/IJISRT24SEP852

Baharin, A. T., Sahadun, N. A., Ramli, S., & Redzuan, N. A. I. (2025). Exploring the adoption of generative artificial intelligence by TVET students: A UTAUT analysis of perceptions, benefits, and implementation challenges. *Journal of Information Systems Engineering and Management, 10*(19s), Article 3052. https://doi.org/10.52783/jisem.v10i19s.3052

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40. https://doi.org/10.3316/QRJ0902027

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Caillaud, E., & Škec, S. (2024). Supervision of design PhD students in an era of generative artificial intelligence. *Proceedings of the Design Society, 4*, Article 47271. https://cnam.hal.science/hal-04766967v1

Chauke, S., Maphosa, C., & Moyo, T. (2024). Postgraduate students' perceptions of the benefits associated with artificial intelligence tools for academic success: The use of the ChatGPT AI tool. *Journal of Curriculum Studies Research, 6*(1), 44–59. https://doi.org/10.46303/jcsr.2024.4

Cowling, M., Crawford, J., Allen, K., & Wehmeyer, M. (2023). Using leadership to leverage ChatGPT and artificial intelligence for undergraduate and postgraduate research supervision. *Australasian Journal of Educational Technology, 39*(4), 89–103. https://doi.org/10.14742/ajet.8598

Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2023). ChatGPT and its impact on research supervision: Insights from Australian postgraduate research students. *Australasian Journal of Educational Technology, 39*(4), 74–89. https://doi.org/10.14742/ajet.8843

Fang, W., Na, M., & Alam, S. S. (2025). Usage intention of AI among academic librarians in China: Extension of the UTAUT model. *Sustainability, 17*(7), 2833. https://doi.org/10.3390/su17072833

Funda, V., & Piderit, R. (2024). A review of the application of artificial intelligence in South African higher education. In *Proceedings of the Conference on Information Communication Technology and Society (ICTAS 2024)* (pp. 44–50). IEEE. https://doi.org/10.1109/ICTAS59620.2024.10507113

Gandhi, R., Parmar, A., Kagathara, J., Lakkad, D., Kakadiya, J., & Murugan, Y. (2024). Bridging the artificial intelligence (AI) divide: Do postgraduate medical students outshine undergraduate medical students in AI readiness? *Cureus, 16*(8), e285501. https://doi.org/10.7759/cureus.285501

Guo, K., Zhang, E. D., Li, D., & Yu, S. (2025). Using AI-supported peer review to enhance feedback literacy: An investigation of students' revision of feedback on peers' essays.

*British Journal of Educational Technology, 56*(4), 1612–1639. https://doi.org/10.1111/bjet.13540

Habibi, A., Muhaimin, M., Danibao, B. K., Wibowo, Y. G., Wahyuni, S., & Octavia, A. (2023). ChatGPT in higher education learning: Acceptance and use. *Computers and Education: Artificial Intelligence, 5*, 100190. https://doi.org/10.1016/j.caeai.2023.100190

Harding, D., & Boyd, P. (2024). *Generative AI and PhD supervision: A covert third wheel or a seat at the table?* UCEM Repository. https://ucem.repository.guildhe.ac.uk/id/eprint/125/1/GenAI_registry.pdf

Hegazy, A. Z., Gaber, S. A., Alkhateeb, I. A., Alqatam, M. A., Almughyirah, S. M., Mahgoub, Y. M., & Shahat, H. A. (2024). Saudi postgraduate students' ethical commitment between awareness and application of artificial intelligence in scientific writing. *International Journal of Learning, Teaching and Educational Research, 23*(10), 583–598. https://doi.org/10.26803/ijlter.23.10.28

Hicham, O., Hayat, E. B., Maryem, E. A., Ouail, C., Mohamed, E.-S., Nacer, M., & Salma, B. (2025). Analysing the adoption of artificial intelligence by Moroccan university teachers: Key insights and implications from the UTAUT model. *Edelweiss Applied Science and Technology, 9*(4), 2722–2732. https://doi.org/10.55214/25768484.v9i4.6644

Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A. A., Hattab, M. K., & Shadid, F. (2024). University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education. *Education Sciences, 14*(10), 1090. https://doi.org/10.3390/educsci14101090

Mbangeleli, N. B. A., & Funda, V. (2024). Mapping the evidence around the use of AI-powered tools in South African universities: A systematic review. In *Proceedings of the International Conference on Education Research.* Academic Conferences and Publishing Ltd. https://papers.academic-conferences.org/index.php/icer/article/download/3180/2837

Mohsin, F. H., Md Isa, N., Ishak, K., & Mohamed Salleh, H. (2024). Navigating the adoption of artificial intelligence in higher education. *International Journal of Business and Technopreneurship, 14*(1), 109–120. https://doi.org/10.58915/ijbt.v14i1.433

Mosae, T., & Kaushal, R. (2025). Navigating the influence of AI tools on students' learning experiences in higher education. *Journal of Informatics Education and Research, 5*(2). https://doi.org/10.52783/jier.v5i2.2627

Ngoc, T. H., Thi, L., Ngan, T., Vu, T., & Uyen, B. (2025). AI tools in learning academic writing: Benefits and challenges for MA students in the English language studies at the Industrial University of Ho Chi Minh City. *International Journal of Artificial Intelligence in Education and Learning, 2*(1), 74–91. https://doi.org/10.54855/ijaile.25215

Nikolic, S., Wentworth, I., Sheridan, L., Moss, S., Duursma, E., Jones, R. A., Ros, M., & Middleton, R. (2024). A systematic literature review of attitudes, intentions and behaviours of teaching academics pertaining to AI and generative AI (GenAI) in higher education: An analysis of GenAI adoption using the UTAUT framework. *Australasian Journal of Educational Technology, 40*(6), 56–75. https://doi.org/10.14742/ajet.9643

O'Leary, Z. (2017). *The essential guide to doing your research project* (3rd ed.). SAGE Publications.

Omar, Z., Harris, M., Harun, M., Irvoni, N., & Ishar, M. (2024). Leveraging artificial intelligence for enhancing postgraduate teaching: A framework for engaging professional learners. *International Journal on E-Learning and Higher Education, 19*(2), 177–196. https://ir.uitm.edu.my/id/eprint/95103/2/95103.pdf

Oubibi, M., Hryshayeva, K., & Huang, R. (2025). Enhancing postgraduate digital academic writing proficiency: The interplay of artificial intelligence tools and ChatGPT. *Interactive Learning Environments*, Advance online publication, 1–19. https://doi.org/10.1080/10494820.2025.2454445

Pashkovskyy, V. M., Pashkovska, N. V., & Tsaryk, I. O. (2023). Perspectives of using artificial intelligence in postgraduate medical education. *SWorldJournal, 20*(2), 61–64. https://doi.org/10.30888/2663-5712.2023-20-02-018

Patterson, A., Frydenberg, M., & Basma, L. (2024). Examining generative artificial intelligence adoption in academia: A UTAUT perspective. *Issues in Information Systems, 25*(3), 238–251. https://doi.org/10.48009/3_iis_2024_119

Rajab, S., Aramadhan, C., & Asad, K. W. (2025). Artificial intelligence adaptation in higher education: Supervisors' efficacy in detecting AI-generated research in postgraduate research. *International Journal of Social Science and Human Research, 8*(9), 7181–7191. https://doi.org/10.47191/ijsshr/v8-i9-59

Ratta, R., Sodhi, J., & Saxena, U. (2025). The relevance of trust in the implementation of AI-driven clinical decision support systems by healthcare professionals: An extended UTAUT model. *The Electronic Journal of Knowledge Management, 23*(1), 47–66. https://doi.org/10.34190/ejkm.23.1.3499

Roy, A. D., Dasgupta, S., Roy, R. Das, Das, D., & Narayan, K. A. (2024). The use of generative artificial intelligence (AI) in teaching and assessment of postgraduate students in pathology and microbiology. *Indian Journal of Microbiology Research, 11*(3), 140–146. https://doi.org/10.18231/j.ijmr.2024.027

Segooa, M. A., Modiba, F. S., & Motjolopane, I. (2025). Generative artificial intelligence tools to augment teaching scientific research in postgraduate education. *South African Journal of Higher Education, 39*(1), 188–205. https://doi.org/10.20853/39-1-6275

Sergeeva, O. V., Zheltukhina, M. R., Shoustikova, T., Tukhvatullina, L. R., Dobrokhotov, D. A., & Kondrashev, S. V. (2025). Understanding higher education students' adoption of generative AI technologies: An empirical investigation using UTAUT2. *Contemporary Educational Technology, 17*(2), ep571. https://doi.org/10.30935/cedtech/16039

Strzelecki, A. (2024). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. *Innovative Higher Education, 49*, 223–245. https://doi.org/10.1007/s10755-023-09686-1

Supianto, A., Widyaningrum, R., Wulandari, F., Zainudin, M., Athiyallah, A., & Rizqa, M. (2024). Exploring the factors affecting ChatGPT acceptance among university students. *Multidisciplinary Science Journal, 6*(12), 2024273. https://doi.org/10.31893/multiscience.2024273

Tarisayi, K. S. (2024). Strategic leadership for responsible artificial intelligence adoption in higher education. *CTE Workshop Proceedings, 11*, 4–14. https://doi.org/10.55056/cte.616

Thong, C. L., Atallah, Z., Islam, S., Lim, W.-L., & Cherukuri, A. K. (2025). AI-powered tools for doctoral supervision in higher education: A systematic review. *Journal of Information & Knowledge Management, 24*(2), Article 2530001. https://doi.org/10.1142/S0219649225300013

Tao, W., Yang, J., & Qu, X. (2024). Utilisation of, perceptions on, and intention to use AI chatbots among medical students in China: National cross-sectional study. *JMIR Medical Education, 10*, e57132. https://doi.org/10.2196/57132

Tian, W., Ge, J., Zhao, Y., & Zheng, X. (2024). AI chatbots in Chinese higher education: Adoption, perception, and influence among graduate students—An integrated analysis utilising UTAUT and ECM models. *Frontiers in Psychology, 15*, 1268549. https://doi.org/10.3389/fpsyg.2024.1268549

University of Pretoria, Faculty of Health Sciences. (2024). *Responsible use of AI for postgraduate students in the Faculty of Health Sciences*. https://www.up.ac.za/faculty-of-health-sciences/news/post_3259711-responsible-use-of-ai-for-postgraduate-students-in-the-faculty-of-health-sciences

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

Wang, C., & Wang, Z. (2025). Investigating L2 writers' critical AI literacy in AI-assisted writing: An APSE model. *Journal of Second Language Writing, 67*, 101187. https://doi.org/10.1016/j.jslw.2025.101187

This edited volume emerged from a pressing and practical question confronting universities worldwide: what constitutes responsible, high-quality postgraduate supervision in an era where generative and analytic AI tools are embedded throughout the research lifecycle? Within a remarkably short period, artificial intelligence has transitioned from peripheral experimentation to an integral aspect of everyday academic practice, influencing how postgraduate candidates conduct literature searches, structure arguments, analyse data, draft chapters, and refine language. However, supervision has never been defined by speed or technical efficiency. At its core, it is an intellectual apprenticeship grounded in mentorship, ethical stewardship, scholarly dialogue, and the development of independent researchers. This book was conceived from the recognition that technological acceleration demands not reactive anxiety, but rather careful, principled reflection on how supervision can remain pedagogically coherent and academically rigorous in the presence of AI.

Ultimately, this book serves as a timely, integrative contribution that synthesises contemporary scholarship and translates it into actionable insights for supervisors, institutions, and policymakers. By addressing ethics, governance, fairness, AI literacy, policy development, and integrity within postgraduate contexts, the volume offers a coherent roadmap for reimagining supervision in a technologically evolving academy. It bridges theoretical discourse and practical application, fostering dialogue between researchers and practitioners while situating AI within broader concerns of equity, quality assurance, and scholarly formation. In doing so, the book not only fills a significant gap in the literature but also provides a strategic foundation for safeguarding the intellectual and ethical standards of postgraduate research amid accelerating digital transformation.