Doktorsavhandlingar från Institutionen för pedagogik och didaktik 89

Stockholm University

# Emerging AI and Ethics in Higher Education

A Technology Mediation Perspective

Alexandra Farazouli

Doctoral Thesis in Education at Stockholm University, Sweden 2026

# Emerging AI and Ethics in Higher Education

## A Technology Mediation Perspective

**Alexandra Farazouli**

Academic dissertation for the Degree of Doctor of Philosophy in Education at Stockholm University to be publicly defended on Friday 8 May 2026 at 10.00 in Lilla hörsalen, Naturhistoriska riksmuset, Frescativägen 40.

**Abstract**

The emergence of AI has become a defining issue for higher education worldwide, and Sweden is no exception. At the same time, emerging AI technologies reconfigure priorities and valuations within educational practice by mediating teaching and learning and opening new paths for how teachers and students relate to higher education practices. In this context, AI-mediated practices raise ethical questions that are often presented as unprecedented yet are deeply rooted in longstanding practices in higher education. This thesis undertakes an empirical exploration of AI-mediated practices in higher education, foregrounding teachers' perspectives and focusing on the ethical issues arising from such mediations. Drawing on postphenomenology and technology mediation theory, the thesis examines how teachers perceive and experience emerging AI artefacts (automated grading systems (AGS) and generative AI (GAI) chatbots) in relation to their practices, and how these artefacts mediate their understandings of what they ought to do and how they ought to act when balancing sometimes competing demands of autonomy and accountability.

The thesis is a compilation of four complementary studies. Study I examines the ethical considerations of AGS, reviewing the literature on AGS and analysing their specificities through a relational ethics approach. This study highlighted that AGS not only introduce technical and procedural considerations but also reconfigure educational practices and relationships in ways that demand ongoing, situated, and relationally attuned ethical reflection. Study II is an interview study with AGS developers who are also university teachers using these systems. It examines their expectations, experiences, and the disruptions that AGS introduce into assessment practices. The findings underscore the ambivalent role of AGS as both promising and disruptive, offering efficiency and consistency, but also introducing 'new' frictions and ethical dilemmas. Study III is a study inspired by the Turing test, followed by focus group interviews with university teachers. It explores how GAI chatbots mediate teachers' perceptions of their assessment practices. The findings indicate that the presence of GAI chatbots, allowing the possibility of AI-generated writing, shapes evaluation practices, prompting teachers to question authorship and, in some cases, reinforcing mistrust within the student–teacher relationship. Study IV is a focus group interview study examining how teachers experience and interpret the emergence of GAI and how it mediates their perceptions of their professional roles. Participants described GAI as both disruptive and potentially transformative. They were compelled to reconsider assessment formats, teaching priorities, and their responsibility to foster critical and ethical engagement with technology.

The combined findings of the four studies show that the emergence of AI unsettles established practices and intensifies the uncertainties that characterise educational situations, placing greater demands on teachers' professional judgment. The thesis also argues that the emergence of AI exposes and amplifies longstanding ethical issues, such as fairness, academic integrity, and equity, reshaping how these issues are understood and enacted as the technologies become embedded in higher education practices.

**Keywords:** *Higher education, emerging technologies, ethics, technology mediation theory, university teachers, artificial intelligence, relational ethics, postphenomenology, automation, generative artificial intelligence, chatbots, automated grading systems.*

Stockholm 2026
http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-253560

ISBN 978-91-8107-552-6
ISBN 978-91-8107-553-3


**Department of Education**

Stockholm University, 106 91 Stockholm

EMERGING AI AND ETHICS IN HIGHER EDUCATION

Alexandra Farazouli


Stockholm
University

# Emerging AI and Ethics in Higher Education

A Technology Mediation Perspective

Alexandra Farazouli

ISBN print 978-91-8107-552-6
ISBN PDF 978-91-8107-553-3

Cover image: Created by the author using an AI image generation tool

Avhandlingar från Institutionen för didaktik och pedagogiskt arbete no 1-10
Doktorsavhandlingar från Pedagogiska institutionen no 1-171

Printed in Sweden by Universitetsservice US-AB, Stockholm 2026

to my grandmother

# Contents

# Acknowledgements

Coming to the end of this doctoral journey, I would like to take the opportunity to express my deepest gratitude to all those who have stood by me with their support, trust, expertise, care, and love. This thesis would not have been possible without you.

First and foremost, to my family – μανούλα, μπαμπάκα και Νίκο. Your unconditional love, unwavering belief in me, and constant encouragement have been my foundation. To my amazing friends - Βελ, Αριάδνη, Ίριδα, Δανάη, Άγγελε, Ανδρέα, Πάνο, Μαρία, Θέμιδα –, to my extended family in Sweden – Κατερινάκι and Ιωακείμ – and of course to my partner in life Hektor, your love and care mean the world to me. Thank you for the endless conversations and analyses, for the good wishes before every conference trip, paper presentation, and submission, and for celebrating every small and big milestone along the way. Thank you for the concerts, the trips, the laughter, the tears, and the moments that kept me grounded, sane, and strong throughout these four years.

There are no words to adequately express how grateful I am to my supervision team, and especially Cormac! Cormac, thank you for being such an exceptional main supervisor – for guiding, inspiring, challenging, and supporting me throughout this journey. Tessy, thank you for introducing me to academic life, for involving me in seminars, workshops, and opportunities, and for your critical and analytical gaze that has strengthened this work at every stage. Klara, thank you for being the best third supervisor I could have hoped for – for your understanding, encouragement, and for believing in me, even during my most uncertain doctoral-student moments.

I would also like to thank my half-time reading seminar readers, Matilda Wiklund and Thomas Hillman, and my final reading seminar readers, Catarina Player-Koro and Gustav Lymer, for their careful reading, thoughtful input, and invaluable advice. Special thanks to Shu-Nu Chang Rundgren for serving as green reader and for helping me polish the final draft of this manuscript.

Returning to the Department of Education (IPD) at Stockholm University after my master's to pursue my doctoral studies has meant more to me than I can express. Teaching, giving seminars, and leading workshops in the same building and classrooms where, five years ago, I took my courses as a newly arrived in Sweden student, has been profoundly meaningful. I would like to thank Åsa Broberg and Matilda Wiklund, who served as Directors of Research

Studies during my doctoral time, for ensuring a safe, supportive, intellectually stimulating, and inspiring working environment. I am also deeply grateful to all my colleagues at IPD.

Within this beautiful IPD circle, I would especially like to thank all the doctoral students for everything we have shared – the support, the conversations, the games, the seminars, the after-works, the fika, the readings, the texts, and the memes. I will miss being a PhD student dearly, and that is very much because of you.

Being part of the Higher Education Learning Practices (HELP) group has been an important developmental opportunity. Thank you to my HELP colleagues for the meetings, retreats, discussions, seminars, and generous exchange of ideas. Thank you to Max for welcoming me into the Swedish Higher Education Research Network (SHERN) community and for the opportunity to receive valuable feedback on two of my articles.

I am also grateful to my colleagues at the Centre for the Advancement of University Teaching (CeUL) for welcoming me into the group – I have learned so much from you. A special thank you to Eva Edman Stålbrandt for leading the digital learning team (DLT), to Andreas Jemstedt and Christophe Premat for our great collaborations in workshops and projects - it has been both an honour and a pleasure to work with you. And once again, Klara, thank you for believing in me.

This work would not have been possible without WASP-HS. It has been an honour to be part of such a vibrant and important graduate school. Beyond funding my doctoral project, WASP-HS provided opportunities to participate in international conferences, supported my semester abroad in Australia, offered tailored courses on interdisciplinary research and AI in the social sciences, and created an invaluable network of researchers across Sweden. I am deeply grateful for the opportunity to have been part of this community. To all the WASP-HS doctoral students – thank you for navigating together what it means to be an AI-curious social scientist and early-career academic.

I would also like to thank the members of the Ethical and Legal Challenges in Relationship to AI-Driven Practices in Higher Education project for sharing this journey of exploring the ethical and legal dimensions of a rapidly evolving field. I learned so much from all of you.

My time in Australia was a formative and inspiring experience near the end of my doctoral studies. I am deeply grateful to my generous hosts. Margaret Bearman at the Centre for Research in Assessment and Digital Learning (CRADLE), Deakin University, and Simon Buckingham Shum at Connected Intelligence Centre (CIC), University of Technology Sydney. Thank you for your trust, inspiration, and for welcoming me so warmly into your research groups. Thomas Corbin, Jack Walton, Juuso Nieminen, Simon Knight, and Antonette Shibani – it was a privilege to be your visiting colleague and to learn alongside such excellent scholars.

Finally, I have been fortunate to meet scholars such as Olga Viberg and Panos Papapetrou, who embody what it means to be not only excellent researchers and passionate teachers, but also deeply human academics. Thank you for setting that example.

As I close this chapter, I do so with profound gratitude. This thesis is not only the result of individual effort, but of collective generosity – of conversations, feedback, encouragement, patience, and love. If this work contributes even a small part to the academic community, it is because I have been surrounded by people who showed me what thoughtful, ethical, and engaged scholarship can be.

Thank you all for walking this journey with me.

# List of studies

**Study I**
Farazouli, A. (2024). Automation and Assessment: Exploring Ethical Issues of Automated Grading Systems from a Relational Ethics Approach. In A. Buch, Y. Lindberg, & T. Cerratto Pargman (Eds.), *Framing Futures in Post-digital Education: Critical Concepts for Data-driven Practices* (pp. 209–226). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-58622-4_12

**Study II**
Figueras, C.*, Farazouli, A.*, Cerratto Pargman, T., McGrath, C., & Rossitto, C. (2025). Promises and breakages of automated grading systems: A qualitative study in computer science education. *Education Inquiry*, *0*(0), 1–22. https://doi.org/10.1080/20004508.2025.2464996

**Study III**
Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2023). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment and Evaluation in Higher Education*, *49*(3), 363–375. https://doi.org/10.1080/02602938.2023.2241676

**Study IV**
Farazouli, A., Cerratto Pargman, T., Bolander Laksov, K., & McGrath, C. (2025). Navigating uncertainty: University teachers' experiences and perceptions of generative artificial intelligence in teaching and learning. *Studies in Higher Education*, *0*(0), 1–16. https://doi.org/10.1080/03075079.2025.2550766

*equal contribution

# List of figures and tables

# Acronyms

| | |
|---|---|
| AGS | automated grading systems |
| AI | artificial intelligence |
| AIED | artificial intelligence in education |
| AIHED | artificial intelligence in higher education |
| GAI | generative artificial intelligence |
| LA | learning analytics |
| TT | Turing test |

# Introduction

From early tools to contemporary algorithms, technological artefacts have been shaping not only what humans do, but how they perceive themselves in relation to the world and their practices (Ihde, 1990; Verbeek, 2005). However, moments of technological upheaval are often experienced or presented as moments of historical rupture. Recently, advances in artificial intelligence (AI) have sparked a significant interest in how these technologies can be used for and integrated into higher education. Although the widespread fascination with AI is often presented as a defining feature of the present moment, such a perception of novelty stems from the situated manner in which technological change is encountered. Technologies become most visible when they disrupt established practices, institutions, and forms of meaning, and it is this disruption that is experienced as radical novelty. What appears new, therefore, is not necessarily the technology itself, but the reconfiguration of relations it introduces – between humans and practices, forms of knowledge, authority, or learning. In this sense, novelty functions less as a property of technological artefacts than as an effect produced by shifts in social and epistemic arrangements, dynamics, and possibilities. Emerging technologies, such as the emerging AI technologies studied in this thesis, reflect and reveal how beings and practices appear within a historical moment (Heidegger, 1977).

As AI currently exerts unprecedented influence on society, it is regarded as the most transformative technology of the twenty-first century (Maslej et al., 2025). Disruptive innovations can transform or create markets, often replacing established practices (Christensen et al., 2015). Similarly, educational technologies have the potential to reshape education by introducing new forms of interaction, governance, and certification (McGrath & Åkerfeldt, 2019). With the rapid advancement of computing and information processing techniques, such as natural language processing and deep learning, AI has been increasingly deployed in educational contexts, giving rise to the field of AI in Education (AIED). Examples of such a deployment include intelligent tutoring systems, learning management systems, and adaptive learning systems (Bond et al., 2024; Chen et al., 2020; Holmes & Tuomi, 2022). Through such examples, AI is becoming entangled with the practices of teaching, learning and assessment within higher education, mediating how these are experienced and enacted.

The emergence of AI artefacts within higher education practices invoke ethical questions that not only demand consideration but may also provoke significant shifts in how institutions respond. Bearman et al. (2020) argue that emerging technologies bring about social and ethical challenges, highlighting that "the interaction of people with technologies creates new ways of perceiving, thinking, doing and relating" (p. 23). This thesis conceives higher education practices not simply as what people do but as nexuses of doings, sayings, and material arrangements that are meaning-making, identity-shaping, and norm-producing activities (Chia & Holt, 2008; Kemmis, 2022; Nicolini, 2009, 2013). From this perspective, ethical issues are not abstract concepts and principles, but are enacted within practices themselves, as technologies participate in shaping what counts as appropriate, legitimate, or desirable conduct. Disruptions introduced by emerging AI can therefore be studied as moments where existing norms and values are unsettled and reconfigured, providing a way to grasp how ethics emerges relationally, through the situated entanglements of people, technologies, and institutions.

One recent example of such a moment, which also comprises a significant part of this doctoral project, is the public release of generative AI (GAI) chatbots, fuelling a vast number of reactions from higher education institutions around the world, manifesting their urgent need for issuing new sets of guidelines on how teachers should deal with new forms of plagiarism from students using GAI chatbots. Such snowballing of reactions, spreading fast among universities around the globe, was also followed by a heated debate on setting the boundaries between the use of GAI chatbots by students as a means for cheating in examinations and as a learning support tool. Thereafter, further questions were raised regarding the appropriateness and suitability of current assessment practices in this new order of things – where GAI chatbots have emerged within educational contexts. Finally, hard-to-answer questions (such as 'what is knowledge?', 'do students own and can claim authorship of GAI chatbots' outputs?', 'what is the purpose of academic assessment?', etc.) were also inspired by the launch of GAI chatbots, potentially mirroring concerns that not only connect with this technology per se, but this technology situated in practice.

AI technologies designed for education are often scripted with a vision to either 'lighten the burden', 'ease administrative workload' or 'inform' teachers and higher education administration to make decisions. This technological mediation, however, introduces ethical questions and challenges related to fairness, responsibility, accountability, privacy and surveillance (Slade & Prinsloo, 2013). Furthermore, technological mediations in educational realities might transform relations among teachers and students, pedagogical acts and administrative procedures. One example of an emerging AI technology provoking such questions and challenges is automated grading systems (AGS), that have been introduced and are being piloted in higher education institutions over the last decade, leading to questions such as: 'who is to be

held responsible when an unfair grade has been assigned to a student?', 'how can the level of influence of such systems on teachers' decision-making for a final grade be measured?', 'what information is being processed by the systems and how?'. Although higher education institutions have been piloting AI artefacts for several decades, there has been limited focus on their ethical aspects (Bond et al., 2024). At the same time, AIED research has repeatedly shown that the introduction of AI into learning environments raises substantial ethical concerns, particularly in relation to privacy, fairness, accountability, transparency, autonomy, agency, and inclusion (Holmes et al., 2022; Sclater, 2016).

Recent AIED literature frequently frames a debate in terms of benefits and risks, or opportunities and challenges for teaching and learning, thereby constructing binary portrayals of the 'positive' and 'negative' effects of AI technologies in higher education practices. Such a dichotomous framing is problematic for several reasons. First, it reduces the discussion to a polarity between 'pros' and 'cons' of AI artefacts, thereby obscuring the nuanced positions that lie between and reinforcing either overly optimistic or overly pessimistic discourses about AI's role in higher education contexts. Second, by conceptualising AI as a monolithic umbrella term encompassing a wide and heterogeneous set of artefacts, it tends to produce abstract and generalised claims about what AI can *do* in higher education, while failing to interrogate the ontological and epistemological dimensions of what AI *is*. Third, conceptualising technology merely as an instrument or tool assumes that it produces direct effects on actors and processes, overlooking its sociotechnical dimensions. These dimensions are articulated through stakeholders' expectations of what AI might offer, their aspirations for AI to address systemic challenges, and their anticipations of its effects on institutional practices and actors (i.e. support, enhancement, replacement scenarios of teachers' work).

In this thesis, I do not attempt to examine the performance of AI in higher education but rather aim to explore how two particular examples of AI artefacts, namely automated grading systems (AGS) and generative AI (GAI) chatbots, shape and are being shaped within higher education practices, drawing on university teachers' experiences and perceptions. I focus on AGS and GAI chatbots, examining how such artefacts participate in, and constitute, meanings and purposes in the context of higher education. I draw on technology mediation theory (Ihde, 1990; Verbeek, 2005, 2011) literature from a vantage point that criticises prominent discourses surrounding AIED literature that focus on the mere instrumentality of technologies and their anticipated implications in higher education. I approach AI technologies as sociotechnical entities (Suchman, 2006; Winner, 1980) that *mediate* and co-shape intentionality and experience of situations. Rather than existing as neutral tools in a vacuum, AI artefacts are co-constituted – interpreted and enacted – through their interactions with people, practices, and contexts; they both shape and are

shaped by how people perceive the world, think, behave, and interact with one another.

Furthermore, I choose to study AGS and GAI chatbots as examples of emerging AI technologies entering higher education. Studying emerging AI technologies is crucial because their roles, functions, and meanings are not yet stabilised, making them key sites for observing how technologies and societies co-shape one another. Unlike established systems, emerging technologies are characterised by uncertainty, novelty, and potential (Rotolo et al., 2015). As Rotolo et al. (2015) point out, the concept of what constitutes *emerging* is challenging to agree on, as it might differ in time, context and the analyst's perspective on what is considered novel. AGS and GAI chatbots could be perceived as emerging AI technologies for higher education stakeholders, such as teachers and students, but not by natural language processing experts and machine learning scientists. More specifically, Rotolo et al. (2015) define emerging technology as:

> a radically novel and relatively fast growing technology characterised by a certain degree of coherence persisting over time and with the potential to exert a considerable impact on the socio-economic domain(s) which is observed in terms of the composition of actors, institutions and patterns of interactions among those, along with the associated knowledge production processes. Its most prominent impact, however, lies in the future and so in the emergence phase is still somewhat uncertain and ambiguous. (p. 1833)

Building on this definition, Pink (2019, 2022) highlights that emerging technologies differ from finished and appropriated products and are experienced as unfinished and unsettled; their meanings and purposes are co-constituted through practices and thereby contribute to shaping how they are embedded in the social sphere.

Furthermore, AGS refer to the technological artefacts designed to evaluate student responses, typically in natural language text form, to predefined questions in digital formats. In contrast, GAI chatbots are designed to generate natural language text responses to questions (prompts), emulating human-like reasoning and dialogues. Although distinct in purpose – one assessing and one producing text – both artefacts are centred in text as the principal medium of interaction and representation, and potentially embody certain kinds of automation within higher education practices. In these terms, AGS and GAI chatbots are not only indicative examples of technological artefacts in their emergence – since they are not considered as established AI artefacts or educational technologies – but also as two sides (stakeholders) of the same coin (practice). For higher education stakeholders, each offers a different route through which reading and writing might be side-stepped or reconfigured; one through the automation of reading and assessing student work, the other through the automation of writing production for reflection, learning and assessment purposes.

At this point, it is important to clarify that, although both technologies are closely associated with assessment practices, they extend to a wider range of educational tasks, activities and practices as well. Accordingly, while this thesis acknowledges assessment as a core educational activity for learning and educational decision-making, it focuses on how AGS and GAI chatbots mediate higher education practices more broadly.

This thesis aims to offer several theoretical and practical contributions. First, it problematises how AI technologies are studied in the current AI and AIED literature, adopting a postphenomenological approach to understanding how technologies and practices are co-constituted through the lens of technology mediation theory. Second, it problematises how ethics is approached in AI and AIED research, arguing for a relational ethics approach to unpack the ethical entanglements of human–technology–mediated experiences of practice in higher education. Third, focusing on university teachers' perspectives, the thesis provides empirical insights into how emerging technologies (AGS and GAI chatbots) are experienced by university teachers, how these technologies shape their perceptions of roles and practices, and the ethical issues that arise when such technologies are introduced into educational practice.

## Aim

This thesis examines how emerging AI technologies mediate higher education practices, focusing on university teachers' perspectives and the ethical issues that emerge through these mediations. Informed by postphenomenology and technology mediation theory, the artefacts in focus are not studied as neutral tools introduced into pre-existing practices, but as active mediators that co-constitute teachers' experiences, perceptions, and actions in higher education. To address this aim, the thesis is guided by three research questions that together unpack the mediating role of emerging AI in higher education. The questions attend to mediation at different but interconnected analytical levels: *experiential*, *perceptual*, and *ethical*.

**RQ 1** How do university teachers experience the emergence of AI in the context of their practices?

This question focuses on teachers' experiences of AI's emergence, foregrounding how AI technologies are encountered and interpreted in relation to their practices.

**RQ 2** How does AI reconfigure university teachers' perceptions of their practices?

While RQ1 addresses the experiential encounter with the emergence of AI, RQ2 examines how these encounters lead to potential reconfigurations of practice. RQ2 focuses on how AI mediates teachers' perceptions of teaching, learning, assessment, and professional responsibility. This question thus examines how AI participates in reshaping how teaching practices are understood and enacted.

**RQ 3** What ethical issues emerge when AI is introduced into higher education practices?

Building on the experiential and perceptual mediations of university teachers explored in RQ1 and RQ2, RQ3 focuses on the ethical issues that arise in relation to AI-mediated higher education practices.

These three research questions are analytically distinct yet conceptually integrated through the technology mediation lens. RQ1 foregrounds teachers' experiential relations with emerging AI technologies, RQ2 examines how these relations reconfigure perceptions and understandings of practice, and RQ3 explores how ethical issues emerge from these AI-mediated configurations. The thesis' research questions enable a postphenomenologically informed analysis of how AI mediates higher education practices and explores the ethical issues that arise from such mediations.

## Outline

The following text offers a framework and synthesis of the doctoral work, which consists of four studies published in peer-reviewed international journals. Together, they address the research questions of the study described in the previous section.

The *Introduction* chapter is followed by the chapter *AI and AI in education,* which introduces the key definitions and terminology related to AI used throughout this thesis. The chapter *AI in higher education* offers a brief overview of previous research in the field of AI in higher education, including studies on university teachers' perceptions of AI in relation to their practices, AGS and GAI chatbots. Next comes the chapter *Emerging AI in the Swedish higher education,* which presents the empirical context of the thesis, focusing on how emerging AI is framed within Sweden's higher education sector, the ways AI has been dealt with by Swedish higher education institutions, and the specificities of university teaching in Sweden. The chapter *Approaching ethics in higher education and AI in education* explores ethical issues related to higher education practice and AIED. Thereafter, the chapter *Theoretical underpinnings* presents the theoretical and conceptual background of the thesis, grounded in postphenomenology and technology mediation theory. The *Methodology* chapter offers an overview of the thesis' overall research strategy and

rationale of methodological choices and concludes with a section on research ethics and researcher positionality. The following chapter includes the *Summaries and findings of the Studies I, II, III and IV*, followed by the *Discussion* chapter, where I discuss how the findings of the Studies I-IV come together and respond to the thesis' research questions, outline the contributions of my research, describe the limitations and delimitations of the thesis and conclude with the implications for practice and research. Finally, in the *Conclusions,* I offer some reflections on the findings and future directions, and conclude the thesis.

# AI and AI in education /*conceptual foundations*

This chapter aims to establish the key definitions and terminology related to AI used throughout this thesis. First, I examine how AI is defined in the literature and the challenges of defining AI. Second, I present a short introduction to AI in education (AIED) as a distinct research field. Finally, I introduce the technologies that are the focus of this thesis: automated grading systems (AGS) and generative AI (GAI) chatbots.

## Defining AI

Approaching the definition of AI is a challenging task. Because of its supra-disciplinary[1] character, there is a lack of consensus among AI researchers regarding a universally accepted definition of AI. Although the term 'AI' has its roots in computer science and engineering, its conceptual development has been significantly influenced by other fields like philosophy, cognitive science, neuroscience, mathematics and linguistics. Each of these fields introduces distinct terminologies, theoretical frameworks, and epistemological assumptions concerning notions such as *intelligence*, *learning*, and *understanding*. For example, in computer science *learning* often refers to the optimisation of algorithms through different methods, such as exposure to data (Murdoch et al., 2019), whereas in cognitive psychology it denotes a change in mental representations or behaviour resulting from experience (Barron et al., 2015). These notions are frequently interpreted in ways that reflect – informed by and constrained by – the disciplinary perspectives used to describe human processes of getting to know, remembering and synthesising information.

Another factor challenging AI researchers to come to an agreement, leading to countless attempts to define AI and determine the boundaries between what is AI and what is not, can be found in the evolving progress of technology and computer science innovations and breakthroughs. According to Holmes and

[1] 'Supra' in this thesis is used to indicate that the complexity of defining AI and AI research involves multiple fields and disciplines in ways that extend across established disciplinary boundaries, including but not limited to inter-, multi-, cross-, and transdisciplinary approaches.

Tuomi (2022), AI definitions depend on the purpose and context of use, while they acknowledge that academia often describes AI from multiple conceptual approaches and disciplines, concluding that "there is no such *thing* as AI" which can be strictly defined as AI. Additionally, the *AI effect,* according to which, terming AI is a moving target across time and technological development, also poses a challenge in defining AI. Systems that are currently called 'AI' might eventually be seen as ordinary systems or software, no longer considered intelligent. This reflects the historical pattern where tasks once seen as intelligent (like speech recognition and autocomplete) stop being labelled as AI once they become commonplace.

Approaching the definition of AI in broader terms, Baker et al. (2019) explain that it does not refer to a single technology but is an umbrella term describing a variety of technologies and algorithmic methods such as machine learning, data mining and neural networks, which can function supervised, semi-supervised and unsupervised, performing certain tasks. Zawacki-Richter et al. (2019), examining how AI is termed in AIED research articles, found that most of the authors do not provide concrete definitions of what AI is, and they rather conceptualise AI as:

> "intelligent computer systems or intelligent agents with human features, such as the ability to memorise knowledge, to perceive and manipulate their environment in a similar way as humans, and to understand human natural language" (p. 10)

AI systems refer to artefacts that can be responsive within an environment and act aimed at optimising their chance of achieving a specific task or objective (Dignum, 2019; Russell et al., 2010). According to Dignum (2019), *AI is not artificial nor intelligence*. Dignum builds this argument based on definitions, highlighting that the 'artificial' only stands there as a counterpart to natural or human intelligence, while 'intelligence' is still an abstract term referring to thinking processes.

Bearman and Ajjawi (2024) identify three main categories of conceptualising and defining AI, which are not mutually exclusive: technical, capability, and relational definitions. Technical definitions focus on describing 'what AI is,' framing AI in terms of the systems' architecture and algorithms. For example, a technical definition might describe AI as a set of algorithms that process input data through layered neural networks to produce predictive outputs. Capability definitions focus on describing 'what AI can do,' defining AI in relation to the roles and tasks it is designed to perform. For instance, a capability definition might describe AI as an image classifier, a speech recognition system, or a driverless car. Finally, relational definitions focus on describing 'how AI works in practice', defining AI as a sociotechnical ensemble of artefacts, humans, and practices. According to Bearman and Ajjawi (2024) a relational definition could emphasise the interactions between humans and AI

artefacts or conceptualise AI as a phenomenon that drives change in professional practice.

## AI in education

Historical roots of AIED can be found in Pressey's (1926) assessment apparatus in the 1920s and Skinner's (1958) teaching machine in the 1950s. However, only recently and in parallel with significant breakthroughs in AI developments, AIED has emerged as a fast-growing field of research and development. AI, often portrayed as the 'new oil' in the public discourse by international organisations (i.e. UNESCO, OECD) and the private sector (i.e. OpenAI, Google), has fuelled a vivid discussion on the tremendous potential and transformative impact of AIED. Advances in AI, along with the growing field of computer-based education, have opened up new applications and possibilities for AI integration into education settings, leading to the establishment of the multidisciplinary term of AIED as an umbrella field for research on AI applications in education, including the field of learning analytics and educational data mining (OECD, 2021). However, as Holmes and Tuomi (2022) point out, although there have been great expectations from AIED developments,

> "[o]ften these expectations have been based on misunderstanding current technical possibilities, lack of knowledge about state-of-the-art AI in education, and exceedingly narrow views on the functions of education in society" (p. 1)

To unpack the complexity of exploring AIED as a field, Holmes (2021) suggests unpacking it into three main areas: AIED as *learning with AI*; AIED as *learning about AI*; and AIED as *preparing for AI*. In the first category, AIED as *learning with AI*, includes research on tools and AI applications that are designed to support and facilitate teaching and learning tasks, such as adaptive learning platforms using machine learning to personalise instruction and feedback (Luckin et al., 2016). AIED as *learning about AI*, mostly refers to the knowledge education that stakeholders need to acquire in order to understand what AI is and how AI works, often described as AI literacy. AIED, as *preparing for AI*, includes aspects of what it is to be living, working and learning in a world where AI technologies are in place, including the impact and challenges that AI bring into practices.

So far, there has been a growing body of AIED literature aiming to map existing and emerging systems and to classify them in categories according to either the use, the system's features or the end user and/or beneficiary. Luckin et al. (2016) focus on three categories of systems – personal tutors, intelligent support for collaborative learning, and intelligent virtual reality. Baker et al. (2019) describe AIED systems as: a) learner-facing – supporting students'

learning; b) teacher-facing – aiding teachers in their teaching (and assessment) activities; and c) system-facing – informing administration at an institutional level. Another classification of systems in higher education, based on the student life cycle by Zawacki-Richter et al. (2019), describes AI systems in education based on their function, falling into four categories: profiling and prediction; adaptive systems and personalisation; assessment and evaluation; and intelligent tutoring systems (see Table 1).

*Table 1 AIED applications in higher education – adapted from Zawacki-Richter et al. (2019, p. 12)*

| **Profiling and prediction** | **Adaptive systems and personalisation** |
|---|---|
| • admissions decisions and course scheduling<br>• drop-out and retention<br>• student models and academic achievement | • teaching course content<br>• recommending personalized content<br>• supporting teachers and learning design<br>• using academic data to monitor and guide students<br>• representation of knowledge in concept maps |
| **Assessment and evaluation** | **Intelligent tutoring systems** |
| • automated grading<br>• feedback<br>• evaluation of student understanding, engagement and academic integrity<br>• evaluation of teaching | • teaching course content<br>• diagnosing strengths and automated feedback<br>• curating learning materials<br>• facilitating collaboration<br>• the teacher's perspective |

Profiling and prediction systems often draw on learning analytics and predictive modelling to identify patterns in student engagement or forecast outcomes such as drop-out risk or course success (e.g. early warning systems). Adaptive systems and personalisation usually adjust content and/or pace to the user's needs based on previous performance or interactions with the systems (e.g. adaptive learning platforms, course recommendation systems). Assessment and evaluation systems include systems designed to automate or assist student work assessment and feedback, as well as teaching evaluation (e.g. automated essay scoring), while intelligent tutoring systems focus on teaching course content and diagnosing learners' strengths and weaknesses. These systems can also curate relevant learning materials, facilitate peer collaboration, and offer insights from the teacher's perspective to support instructional decision-making.

Finally, Holmes and Tuomi (2022) also classify AIED systems into three main categories according to the role of the end-user and main beneficiary:

student-focused, institution/system-focused and teacher-focused. Table 2 presents an overview of the taxonomy of AIED developed by (Holmes & Tuomi, 2022) including several examples of already researched or available AIED applications.

*Table 2 Taxonomy of AIED - adapted from Holmes and Tuomi (2022, p. 550)*

| **student-focused** | **teacher-focused** |
|---|---|
| • Intelligent Tutoring Systems (ITS)<br>• AI-assisted Apps (e.g., maths, text-to-speech)<br>• AI-assisted Simulations (e.g.,VR, AR)<br>• AI to Support Learners with Disabilities<br>• Automatic Essay Writing (AEW)<br>• Chatbots<br>• Automatic Formative Assessment (AFA)<br>• Learning Network Orchestrators<br>• Dialogue-based Tutoring Systems (DBTS)<br>• Exploratory Learning Environments (ELE) | • Plagiarism detection<br>• Smart Curation of Learning Materials<br>• Classroom Monitoring<br>• Automatic Summative Assessment<br>• AI Teaching Assistant (including assessment assistant)<br>• Classroom Orchestration<br>**system-focused**<br>• Admissions (e.g., student selection)<br>• Course-planning, Scheduling, Timetabling<br>• School Security<br>• Identifying Dropouts and Students at risk<br>• e- Proctoring |

Such a taxonomy highlights the breadth of AIED across the higher education ecosystem. It shows that AI does not operate as a single pedagogical tool but rather as a network of systems designed and purposed for different stakeholders – students, teachers, and institutions. From a pedagogical perspective, this categorisation underscores the importance of understanding how these technologies interact and potentially reshape roles and relationships within higher education. While student-focused systems may directly influence learning processes, teacher- and system-focused systems raise questions about agency, data ethics, and institutional responsibility. In this sense, Holmes and Tuomi's taxonomy provides a useful lens for examining both the functional diversity and the pedagogical implications of the introduction of AIED systems.

At this point, I would like to clarify that in this thesis, I focus on automated grading systems (AGS) that are designed and purposed for teachers in higher education, and generative AI (GAI) chatbots which are considered as a general-purpose AI technology. AGS are typically classified as AIED technologies because they are explicitly designed to perform educational tasks – evaluating student work, providing feedback, or supporting assessment processes – within pedagogical frameworks and institutional settings. In contrast, general-purpose GAI chatbots are not, in themselves, considered AIED systems. Although they can and tend to be used for educational purposes, they were not developed taking into account instructional design principles, learner modelling, or educational objectives. Their educational role, therefore, depends on how they are integrated and adapted by teachers, researchers, or institutions.

## Automated grading systems

Automated Grading Systems (AGS) are technological artefacts designed to evaluate student responses to pre-defined assessment tasks, particularly those involving written natural language input. These systems are used to score a wide range of question formats, including multiple-choice and true/false questions, as well as more complex formats such as short-answer and essay-based responses. Automated grading tools for assessing multiple-choice and true/false questions have long been embedded as available functions in most of Learning Management systems (LMS) – like Canvas, Blackboard, It's learning, Google Classroom, etc. – and have the capacity to automatically assign scores in such questions to facilitate and speed up assessment. AGS enforced by machine learning and natural language processing methods include Automated Essay Scoring (AES), which typically focus on features such as linguistic style, coherence, and structural organisation (Balfour, 2013; Larkey, 1998), and Automated Short-Answer Grading (ASAG) which aim to determine the semantic accuracy, completeness, and relevance of student responses in relation to the assessment prompt (Brooks et al., 2014; Leacock & Chodorow, 2003; Pulman & Sukkarieh, 2005).

## Generative AI chatbots

Generative AI (GAI) chatbots are designed to generate natural language responses to questions (prompts) in digital formats and can be classified as an example of generative AI. GAI refers to a category of AI based on machine learning models pre-trained on large datasets that can produce new content such as text, images, audio, or other data forms. Large language models (LLMs) are a class of artificial neural networks designed for the processing and generation of natural language text. Through deep learning techniques applied to extensive textual corpora, these models identify linguistic patterns and structures that enable them to produce new text in response to given prompts or inputs (Jurafsky & Martin, 2023).

GAI chatbots' history dates back to the 60s with MIT's ELIZA. Major advances in natural language processing have occurred since the introduction of the transformer architecture in 2017 (Vaswani et al., 2017). This architectural shift enabled LLMs to handle extensive textual datasets more efficiently and to perform a wide range of language-related tasks with improved effectiveness. Building on this development, OpenAI subsequently released the first Generative Pretrained Transformer (GPT) model, trained on large-scale text corpora and capable of producing text across multiple styles and genres. Later in 2019, GPT-2 (Radford et al., 2019) was released as an even larger and more powerful model, with 1.5 billion parameters and the ability to generate more quality text in terms of coherence and human-like text on a wide range of

topics. This model was followed by GPT-3 (Brown et al., 2020) in 2020, including 175 billion parameters, and introducing the concept of few-shot learning, in which the model could be trained to perform new tasks with just a few examples of labelled data.

The first commercial GAI chatbot, ChatGPT, was initially launched as a variant of GPT-3 specifically designed for conversational applications. It was also developed by OpenAI and uses the same transformer architecture as GPT-3, while updated versions in early 2023 evolved to GPT3.5 and GPT4 (OpenAI, 2023). However, ChatGPT is trained on a specific dataset of conversational data and is fine-tuned to generate responses that are more natural and relevant to a conversation. Except for OpenAI's ChatGPT, there have been more GAI chatbots monopolising the interest of the educational research community, such as Google's Gemini[2] (Thoppilan et al., 2022) and BigScience Bloom (Scao et al., 2023).

Concluding this chapter, in this thesis, to address the challenge of approaching AI in too broad terms, I focus on two specific groups of artefacts – AGS and GAI chatbots. Aligning with Bearman and Ajjawi (2024), I approach AI in relational terms by examining how these artefacts mediate higher education practices, with particular attention to university teachers' perspectives and the ethical issues that emerge through these mediations. With regard to the choice of these two technologies, they are examined as examples of emerging AI in higher education. The inclusion of both AGS and GAI chatbots serves to capture the diversity of ways in which AI is enacted and mediates higher education practices. While AGS are purposed by design for higher education practices, GAI chatbots are purposed in practice by higher education stakeholders, teachers and students. This contrast in the artefacts provided multiple empirical entry points to explore how AGS and GAI chatbots mediate teachers' practices and the ways in which their functions are co-constituted by teachers within those practices. Additionally, since we are on the brink of technological expansion in the field of natural language processing and large language models, this thesis aims to contribute to the pressing need for empirical research examining both AGS and GAI chatbots in context and placed in situations where insights can be drawn about the impact they have on the living world and key sectors, such as in education.

[2] Gemini was previously named Bard

# AI in higher education */previous research*

This chapter aims to present a brief overview of previous research in the field of AI in higher education, including studies on university teachers' perceptions of AI in relation to their practices, AGS and GAI chatbots.

Research on AI in higher education (AIHED) has grown significantly in recent years, yet remains an emergent and uneven field in terms of scope and focus. Since 2021, publication rates on research examining AI in higher education have almost tripled compared to earlier years (Crompton & Burke, 2023). Much of the early work was mainly led by computer scientists and data scientists (Chen et al., 2020; Zawacki-Richter et al., 2019), often adopting a techno-romantic view of AI developments, representing techno-enthusiastic readings of AI decontextualised from practice or purpose potential, and adopting largely uncritical narratives of AI integration in higher education institutions. Only recently have researchers from education departments begun contributing in greater numbers, signalling a gradual broadening of perspectives (Crompton & Burke, 2023). However, despite this surge of research interest, the current body of AIHED research remains limited in several significant ways. Empirical studies are still scarce, and much of the existing literature centres on students' uses of AI rather than teachers' experiences and practices (Crompton & Burke, 2023; Zawacki-Richter et al., 2019). Additionally, the vast majority of studies do not engage with learning theories or theoretical frameworks and often rely on speculative or hype-inflated claims (McGrath et al., 2025). At the same time, various higher education institutions have been using and piloting AI-driven approaches such as data mining and machine learning in order to support and enhance educational activities (Beseiso et al., 2021; Forsberg et al., 2015; Gierl et al., 2014), while the broader field of AIED is increasingly shaped by commercial interests, with 30 EdTech unicorns valued at over 89 billion USD as of January 2023 (GMI, 2023; Holmes et al., 2022). This accelerating institutional and commercial uptake contrasts sharply with the still-fragmented research landscape, underscoring the need for more critical, theoretically informed, and empirically grounded studies.

Reviewing research on AIHED, Zawacki-Richter et al. (2019) found that the majority of empirical studies had followed a quantitative research design, while descriptive and pilot studies on certain AIED systems were mostly conducted in STEM and computer science departments representing a more technology-oriented and less pedagogy-focused point of view. Hence, the same

authors highlight that there is a lack of qualitative studies exploring AIHED integration in education practices. However, Zawacki-Richter et al. (2019) argue that the AI applications in higher education "provide enormous pedagogical opportunities for the design of intelligent student support systems, and for scaffolding student learning in adaptive and personalised learning environments" (p. 21). In addition, the authors point out that such systems might be of great benefit, especially for mass higher education and large higher education institutions which offer open and distance courses, as AIHED could increase access to a vast number of students. Among the benefits of AI in higher education highlighted in the review, Zawacki-Richter et al. (2019) claim that AI can assist in providing flexible, interactive, and personalised learning opportunities, by "relieving teachers from burdens, such as grading hundreds or even thousands of assignments, so that they can focus on their real task: empathic human teaching" (p. 21). Finally, turning their focus on the ethical and pedagogical implications of AI integration into higher education institutions' practices, Zawacki-Richter et al. (2019) note that there has been an absence of critical reflections found in the AIHED research about the potential ethical risks and pedagogical impacts of such systems. This could be potentially explained by a) the 'atheoretical syndrome' occurring in the field of educational technology, which often misses explicit engagement with theory (Hew et al., 2019); and b) the limited representation of researchers affiliated with Education departments (Zawacki-Richter et al., 2019).

Several discourses on AI in higher education have been identified by education scholars, reflecting how AI is portrayed and how it impacts higher education institutions, education stakeholders and educational activities. Selwyn (2016), discussing the transformative power of digital technologies in education settings, identifies a range of discourses on how digital technologies mediate non-digital processes and practices in education, introducing several new possibilities and limitations. According to Selwyn (2016), the role of digital technologies leading to educational changes ranges from *improvements* to *revolutions* (see Table 3). Selwyn (2016) makes a call for careful and critical thinking about education and the digital, stating that "The ideas of digital improvement/transformation/disruption of education clearly require problematising: that is, taking a step back from them and not taking them at face value" (p. 23). More specifically, Selwyn (2016) pinpoints that: a) claims that portray digital technologies as 'game-changers' aiming to 'fix' education, need to be examined in relation to the wider agendas, beliefs and interests for education reform they are linked to; b) corporate, commercial and economically driven claims for the deinstitutionalisation of education are aligned with business ideals, corporate interests and commercial values and not education or pedagogical values solely; and c) 'technical fixes' are rarely adequate for addressing societal (and therefore educational) challenges.

Selwyn (2016) cautions that while digital technologies may offer promising solutions to educational challenges, they also bring about unforeseen outcomes and changes that might not be initially considered. For instance, introducing a new digital learning platform might improve access to educational resources and facilitate collaborative learning. However, it might also inadvertently widen the digital divide, create privacy issues, or lead to over-reliance on technology. Selwyn (2016) also refers to 'longer-term shifts' as potential changes in teaching methods, student behaviours, or even the entire educational system as a result of the integration of digital solutions.

*Table 3 Digital Technology and Educational Change (Selwyn, 2016, p. 5-8).*

| Educational Change | Description | Example |
|---|---|---|
| *Improvement* | *Technology enhances/supports/ enables teaching and learning hence education is improved and upgraded but remains essentially the same* | expanding the capacities of teachers |
| *Transformation* | *Technology 'hacks' and 'reboots' educational processes and practices transforming education by acting as a facilitator of new forms of teaching and learning* | online game-based learning |
| *Revolution* | *Technology challenges the status quo, redistributes power and destabilizes 'education establishments' driving a radical change of the forms and purpose of education* | advancing the interests of private markets over public sector monopolies |

Along the same line, Bearman et al. (2022), motivated by the vague and debatable references to AI found in the literature in the past years, conducted a critical literature review examining how AI is invoked in higher education research.

> "How universities respond to AI depends not only on what AI is but also on what it is *understood* to be. The ways that AI is portrayed within the higher education literature helps shape research, policy and practice, and discourses about technology can be powerful." (Bearman et al., 2022, p. 370)

Drawing on Fischer (2010), Bearman et al. (2022) argue that discourses about the centrality of technology influence how integral it becomes in our lives and society. Among their findings, Bearman et al. (2022), found that only a few articles included definitions of AI. However, two discourses about AI have emerged: a discourse of *imperative response,* revolving around the emergence of significant socio-technical shifts, highlighting the urgent need for higher education to react; and a discourse of *altering authority,* focusing on how AI

is reshaping the authority and agency dynamics within academia. Bearman et al. (2022) also found a prevalent dualism of *dystopia-is-now* and *utopia-just-around-the-corner* throughout these two discourses. More specifically, within the discourse of imperative response Bearman et al. (2022) identified that higher education institutions either need to resist AI changing force or are required to respond positively to AI transformation. The discourse of altering authority highlights the shifting role and purpose of teachers, as AI assumes a more prominent role, as well as the in ways which students can exist and function in an increasingly AI-powered educational reality. The dipole of present dystopia and near-future utopia is present in this discourse too, implying a potential loss of agency for the former and beneficial enhancement in the latter.

Historically, educational technology has repeatedly been positioned as a catalyst for change, presented through narratives of disruption and renewal that promise to reshape established pedagogical practices. One illustrative example of such technology can be found in the emergence of massive open online courses (MOOCs). MOOCs were widely celebrated as a disruptive force expected to democratise access and transform higher education (McGrath et al., 2017; Ross et al., 2014; Yuan & Powell, 2013). Despite this early enthusiasm, their anticipated revolutionary impact did not fully materialise, as MOOCs largely served already well-educated professionals rather than broadening participation among new learner groups (Eisenberg & Fischer, 2014; Stöhr et al., 2019). In other words, although MOOCs did not fulfil the expectations initially attached to them, they came to serve different purposes, meeting the needs of distinct learner populations. Similarly, earlier technologies such as interactive whiteboards (IWBs) were introduced under the promise of pedagogical innovation but were often used in ways consistent with existing teaching practices (Moss & Jewitt, 2010). These patterns suggest that technological change in education is frequently shaped by existing pedagogical and institutional contexts, where both stakeholders and technologies find their purpose situated within relational educational realities rather than radically transforming them. In this light, the emergence of AI in higher education can be seen as part of this broader historical continuum. Like earlier educational technologies, AI is accompanied by powerful narratives of transformation – promising personalised learning, administrative efficiency, and new modes of assessment. The focus of this thesis, therefore, lies not in evaluating what AI can do for higher education but in examining how AI mediations reconfigure and reshape academic practices, professional roles, and the values that underpin higher education itself.

Finally, as Bond et al. (2024) indicate, there is a research gap in AIHED research engaging with ethical and contextual considerations of studies exploring AI integration in higher education institutions. More specifically, the authors conclude that ethical implications of AI deployment in higher education institutions have emerged as a topic that requires further attention from

researchers and identify that "[t]here is a loud and resounding call for an enhanced focus on ethics in future AIHED research, with 40.9% of reviews in this corpus indicating that some form of ethical considerations are needed." (Bond et al., 2024, p. 34). In addition, further empirical research is needed as pointed out by several studies, examining specific AI artefacts in higher education settings and focusing on teachers' perspectives (Alkhalil et al., 2021; Buchanan et al., 2021; Bond et al., 2024).

In response to these identified gaps, this thesis undertakes an empirical exploration of AI-mediated practices in higher education, foregrounding teachers' perspectives, and focusing on the ethical issues arising from such mediations.

## Studies on AI and university teachers

Currently, research on university teachers' perceptions of AI in higher education is limited and fragmented. Existing studies have primarily examined teachers' motivation to adopt AI, their willingness to use AI, and their perceptions of how AI may be integrated into higher education practices (Herodotou et al., 2021; Nguyen, 2023). However, this body of literature reports inconsistent and sometimes contradictory findings, reflecting considerable ambiguity about the benefits and challenges of AI in higher education (Lee et al., 2024; Nguyen, 2023). In line with this ambivalence, research specifically addressing GAI reveals mixed responses by teachers. Dakakni and Safa (2023) found that a substantial proportion of teachers in their study express distrust toward GAI chatbots, largely due to concerns about student plagiarism. Nevertheless, most participants in that work supported training in GAI, particularly as a means of monitoring and managing academic integrity.

Several studies suggest that teachers' acceptance of AI is shaped by how its use is motivated and supported institutionally. For example, McGrath et al. (2023), examining the conditions under which teachers perceive AI as legitimate or appropriate, found that teachers are more receptive when AI is framed as a tool for promoting equity or supporting students, such as first-generation learners or students with disabilities. At the same time, teachers frequently report limited understanding of AI systems and uncertainty about institutional policies and expectations surrounding their use (McGrath et al., 2023; Nguyen, 2023). Complementing these findings, Kohnke et al. (2023) argue that teachers' familiarity with and confidence in understanding and using AI play a crucial role in adoption, while also emphasising the need for targeted professional development to address the specific challenges faced by language educators.

Beyond attitudes and adoption intentions, qualitative research has begun to explore how AI technologies may reshape teachers' professional values and practices. Sporrong et al. (2025) examine university teachers' anticipations of

AI in assessment and show that teachers view AI as simultaneously enabling and constraining their ability to act in accordance with professional, pedagogical, and relational values. Rather than emphasising tool use, their study foregrounds teachers' valuing practices, highlighting tensions related to automation, student-teacher relationships, and professional judgment. Overall, this limited body of literature reveals existing tensions between perceived opportunities, challenges and professional uncertainties surrounding AI use in higher education.

## Studies on automated grading systems

Automated grading systems (AGS), such as automated essay scoring (AES) and automated short answer grading systems (ASAG), are designed to grade free-text responses written in natural language where students might need to answer questions, provide definitions, offer explanations, or analyse and discuss various topics. Although these types of questions offer significant benefits in terms of learning since students are asked to recall external knowledge and elaborate on a specific topic (McDaniel et al., 2007), those student responses could contain several ambiguities, negations and similarities that may require extra attention from educators as well as being influenced by biases related to e.g. culture, writing style and language proficiency. Additionally, in regard to human grading fairness and bias due to grader subjectivity or tiredness, AGS can potentially assist teaching staff with feedback and explanations of the scores that the system suggests (Süzen et al., 2020).

In this sense, research on AGS has primarily presented them as support systems for teachers and focused on their capacity to manage teachers' workloads by reducing the burden of grading large amounts of complex assignments and exams. Such narratives often argue that by using AGS in their practices, teachers can allocate a greater proportion of time to other pedagogical tasks, such as mentoring and coaching students. Furthermore, AGS systems are said to extend educators' capacities, assisting their grading decisions by providing valuable insights, revealing patterns of common mistakes and identifying exceptional student responses. Last but not least, another commonly cited benefit concerns timeliness, as system-assisted grading can significantly reduce turnaround times for feedback and grades (Sung et al., 2019).

Notwithstanding the promised benefits of AGS, there are several ethical concerns about them related to the accuracy of grades' predictions, the pedagogical choices employed by these systems, issues of fairness, accountability and transparency and the level of influence on teachers' decision making (Holmes et al., 2022). Among some limitations of such systems, Gierl et al. (2014) have shown that AES systems may not be accurate and effective in all writing genres and small classes due to limited available data of previous graded assessments. Furthermore, the quality of extensive feedback provided

by such systems (Dikli, 2010) has been reported to be relatively low, as the more detailed the feedback, the more questions students were asked about their assessment decision (Barker, 2011).

Forsberg et al. (2015) in their study of evaluating a scoring and semi-automated grading model in postgraduate nurse education, report that teachers appreciated this model because of improved reliability of exam results and decreased workload. The data provided by this semi-AGS to teachers facilitated the correction of the exams, and students could receive feedback much more quickly. However, it was also mentioned that the model still needs further improvements since higher grades were wrongfully assigned to some cases, and corrections to the system by teachers were significantly time-consuming. At this point, it is important to mention the risk of teachers not perceiving AGS as a relative source of guidance and support but unquestioningly trusting them (Selwyn, 2016). This approach is problematic for two reasons: first, because education and pedagogical activities should be seen in their whole and recognise their complexity as they cannot be limited solely to data analysis and algorithms. Second, because in this way, the experience of teachers and quality of teaching might be lowered due to over-trusting such tools.

In line with this research, Yang and Gulbahar's (2025) study on the use of AGS for assessing programming assignments argues that AGS "cannot replace human judgment, it can efficiently perform an initial assessment of coding assignments" (p. 38). This position promotes a hybrid, semi-automated model of human-AGS collaboration, in which AGS supports teachers by providing preliminary evaluations or rubric-based feedback that teachers then subsequently review and finalise. However, the literature remains limited in addressing the epistemic implications of such collaboration – specifically, the extent to which teachers may come to rely on, defer to, or uncritically accept AGS outputs as authoritative during the grading process. Moreover, the notion of 'semi-automation' itself deserves further examination. One could argue that automation is not a matter of degree but a categorical condition: a system either automates a task or it does not. From this perspective, the term 'semi-automated grading' risks obscuring the actual locus of decision-making authority and accountability within the assessment process.

## Studies on generative AI chatbots

Research on AI in education and technology-enhanced learning has increasingly focused on chatbots (Bahja et al., 2019), particularly in relation to self-regulated learning processes (Maldonado-Mahauad et al., 2022) and applications in intelligent tutoring systems (Mirzababaei & Pammer-Schindler, 2022). Prior research on AI chatbots highlights availability and perceived usefulness as key factors contributing to their success, while identifying limitations such as low-quality content, insufficient regulatory frameworks, and

weak conversational design as primary causes of failure in practice (Janssen et al., 2021; Kasneci et al., 2023). Additionally, another stream of publications engages with the revolutionary potential of GAI chatbots to transform education and the strategies education stakeholders need to develop to tackle the risks of possible harms and disruptions in the context of education (Gilliard & Rorabaugh, 2023; Saunders, 2023; Tlili et al., 2023).

Studies, using the Turing-Test[3] experiment design, have explored the capabilities of GPT3 in writing student essays, by finetuning the model and engineering prompts (Choi et al., 2023), discussing the capabilities and flaws of GPT3 when manipulating its parameters, the prompts, reporting that although the quality of the essays could be high and often imaginative and innovative, there is still a lot of room for development (Bommarito & Katz, 2022; Choi et al., 2023; Elkins & Chun, 2020). Such studies suggest that GAI chatbots, based on GPT models and similar technologies, could support students in their creative writing and potentially be used as 'commonplace' tools such as grammar checkers.

In the past years, several studies have been conducted aiming to explore GAI chatbots' capacity in responding to exam prompts in higher education. One study examining how ChatGPT would perform in law school exams by testing its answers to four real exams reported that ChatGPT responses passed the exam with an average grade of C+[4] and could potentially be beneficial to low-performing law students to improve their writing and improve their grades in exams. Another study has tested ChatGPT in several open-ended exam questions from an MBA course in Operations Management, where the chatbot response would be evaluated as B to B-. Additionally, a number of studies (Generative Pretrained Transformer et al., 2022; Zhai, 2022) have also inquired whether GPT3 could author an academic paper with the minimum human involvement, suggesting that GPT3 could be a useful tool for academic writing.

Finally, while existing studies on GAI chatbots in higher education discuss and/or problematise their potential to support teaching and learning practices, important research gaps remain. So far, literature focuses on short-term evaluations, simulations of integrating GAI chatbots into pedagogical activities, or performance testing within isolated courses, leaving a gap of qualitative research situated in teachers' practices, as well as longitudinal, large-scale, and cross-disciplinary studies.

[3] The Turing Test was originally introduced by Alan Turing as the 'imitation game', a thought experiment designed to operationalise the question of whether machines can be said to think (Turing, 1950).

[4] In this article the authors use an A to F grading scale, in which C+ represents an above the average grade.

# Emerging AI in Swedish higher education */empirical context*

This chapter provides the empirical context of the thesis. It examines how emerging AI is framed within Sweden's higher education sector, the ways AI has been dealt with by Swedish higher education institutions, and finally concludes with a section on university teaching in Swedish higher education, that briefly presents the specificities of the role of the university teachers in the Swedish context.

Emerging technologies have become a focal point of scholarly inquiry and a prominent subject of policy discussions. Such an increasing interest is evidenced by the growing volume of academic publications and media coverage referencing emerging technologies as a phenomenon of its time (Rotolo et al., 2015). Unlike established technologies, emerging technologies are characterised by uncertainty, rapid development and disruptive potential, often opening new avenues of scientific inquiry while also posing novel ethical, legal and regulatory challenges. Studying emerging technologies as unsettled artefacts that are not yet fixed in terms of their design, development and deployment for and within practices, is of importance as it enables critical reflection on how they come to participate in practices and opens the space for exploring how they become embedded within them.

In recent years, the emergence of AI has become a defining issue for higher education around the world, and Sweden is no exception. AI has been portrayed by international organisations such as UNESCO and OECD as a technology with the potential to transform core academic practices and a key driver of future educational innovation. More specifically, UNESCO claims that AI is capable to "address some of the biggest challenges in education today, innovate teaching and learning practices, and ultimately accelerate progress towards SDG 4" (Miao et al., 2021, p. 3), while OECD describes AI as an inevitable driver of innovation and change (OECD, 2023); framings that emphasise both promise and urgency. Echoing such approaches, at a national level, Swedish policy discourse increasingly positions AI as a strategic priority for maintaining the country's competitiveness and fostering innovation. The national AI commission's (2024) roadmap for Sweden can be considered as a recent example of such discourse, that opens with a foreword stating that:

> "AI is a disruptive, emerging technology that, in human hands, creates the conditions for fundamental improvements to our societies, just as the railway, electricity and the telephone did. (…) AI becomes a threat if we stand on the sidelines and passively watch a technological and social change without acting and linking it to our overall endeavours in Sweden for a better life for all." (p. 3)

In this document, the AI commission describes AI as a technological development that Sweden must necessarily embrace. Such a narrative (or *sociotechnical imaginary*[5]) about AI as not merely a tool but a national imperative expresses a vision for a desirable technological future to mobilise political and institutional action.

While public debate has emphasised the need for Sweden to "keep up" in the global AI race for innovation, this framing has often overlooked the complex ethical, and legal challenges that higher education institutions face. Within the Swedish higher education sector, perspectives on AI have been more unsettled, highlighting the challenges and disruptions that AI introduces to educational practices, as well as the blurred boundaries between acceptable and unacceptable uses of GAI (Erhardt et al., 2025; Sporrong, McGrath, Viberg, et al., 2025). This has been particularly evident in the field of assessment, where the emergence of GAI artefacts has raised concerns about academic integrity and the validity of assessment methods (Kyrk et al., 2023). The launch of GAI chatbots in particular has raised concerns about academic misconduct and has been perceived as opening the possibilities for cheating (Erhardt et al., 2025; Sporrong, McGrath, Viberg, et al., 2025). Reflecting these concerns, in 2023 and 2024, the Swedish higher education authority (UKÄ)[6] noted that a sharp increase has been reported in the number of suspensions and warnings to students for unauthorised use of AI (Kyrk et al., 2024). Early responses from universities were often reactive and restrictive, such as attempts to ban the use of AI and consider AI use by students as committing acts of academic dishonesty (often also referred to as AI-plagiarism).

According to UKÄ, in 2023, over a third of Swedish universities had introduced new measures, such as developing guidelines, training, and revised examination formats, to address the risks of students using AI to cheat or mislead during exams (Bendixen et al., 2024; Kyrk et al., 2023; Premat & Farazouli,

---

[5] *Sociotechnical imaginaries* can be understood as shared visions of how society should be organised and function, which are embedded in and pursued through nation-specific scientific or technological initiatives (Jasanoff & Kim, 2009, 2015; Sismondo, 2020). These imaginaries are closely linked to the exercise of state power, shaping actions such as setting development priorities, distributing resources, building infrastructures, and directing political debate or dissent.

[6] UKÄ stands for Universitetskanslersämbetet and is the government agency responsible for evaluating the quality higher education and research. It produces official statistics and monitors compliance with laws and regulations among universities and university colleges.

2025). As an attempt to support higher education institutions in their efforts to introduce such measures, UKÄ published in 2023 a report named *Artificiell intelligens och högskolans utbildningsutbud* (Artificial Intelligence and the Range of Higher Education Programmes). In this report, focusing on the realisation of the potential and economic value of AI in the public sector, in terms of higher productivity and increased revenues, leading to the urgent national demand for developing AI expertise, UKÄ recognised the crucial role of higher education institutions to capitalise on AI potential (Odelberg et al., 2023; Universitetskanslersämbetet, 2024). Therefore, currently, the focus is on how universities can develop foundational AI competence and expertise with and within courses and programs.

While Swedish policy debates have so far focused mainly on students' use of AI, the use of AI by teachers, although encouraged (Sporrong, McGrath, Viberg, et al., 2025), remains underexplored. UKÄ has provided little discussion and guidance regarding how university teachers may or may not use AI in their practices, leaving gaps about the legal and ethical framing of teachers' professional practice. This gap becomes especially significant in relation to the Swedish legal and ethical framework surrounding the concept of *myndighetsutövning* (exercise of public official authority) and the assessment of students' work. Until 2017, grading was explicitly described as a form of myndighetsutövning, emphasising that teachers act as representatives of the state and exercise formal power over students when awarding grades (Herjevik, 2017). Although a 2020 revision removed the term from formal regulatory language (Herjevik, 2020), many universities – including Uppsala University[7], Stockholm University[8], and Örebro University[9] – continue to describe grading in these terms, stressing teachers' accountability when making assessment decisions.

AI systems that are designed to automate (fully or partly) teachers' tasks, such as grading, may undermine the notion of the teacher as the accountable decision-maker. Additionally, merely keeping teachers "in the loop" may not be sufficient to satisfy the Swedish legal and ethical expectations attached to myndighetsutövning, since the concept does not only require that teachers have the *possibility* to intervene, but that they actively exercise educational

[7] https://www.uu.se/download/18.3691d73a192907d48ab311f3/1729680999217/R%C3%A4ttss%C3%A4ker%20examination%202023Tillg%C3%A4nglig.pdf (Retrieved February 26, 2026)

[8] https://www.su.se/enheter/juridiska-institutionen/nyheter/nyhetsartiklar/2022-12-12-reviderade-regler-for-examination-fr.o.m.-vt23 (Retrieved February 26, 2026)

[9] https://www.oru.se/utbildning/jag-ar-student/mina-studier/studenters-rattigheter-och-skyldigheter/regler-for-examination/examinator-roll-och-ansvar2/ (Retrieved February 26, 2026)

judgments and are fully accountable for the decisions made. If AI systems perform the substantive evaluative work while teachers only validate AI systems' outputs, the teacher's role risks becoming procedural rather than judgment-based (Colonna, 2024). In such cases, formal responsibility might remain with the teacher, but de facto decision-making power would lie with the AI system, creating a mismatch between legal accountability and practical control. This raises profound questions about what level of human oversight constitutes genuine decision-making, and whether delegating core evaluative functions to AI could erode the integrity of grading as an act of myndighetsutövning. These concerns highlight that the debate on AI in Swedish higher education institutions cannot be limited to questions of technological efficiency, organisational productivity or control of student uses of AI, but requires attention regarding teachers' professional practice.

## University teaching in Swedish higher education

Swedish higher education operates within a governance framework characterised by state steering combined with extensive institutional responsibility. At the same time, it is underpinned by firm commitments to academic freedom and professional autonomy, which relate not only to research but also explicitly to teaching (Ekberg, 2025; Ekberg & Söderbergh Widding, 2025). Considering that Swedish higher education institutions function as public authorities within a system of national regulation and accountability is central for understanding the role of the university teacher, whose autonomy is formally recognised yet structurally situated within national legislation and institutional frameworks.

The role of university teachers in Sweden is generally understood to retain substantial control over the form and content of teaching. As Carlson (2016) notes, "(t)eachers are not only to freely conduct research, but also to freely form their teaching and teaching materials" (p. 362). This understanding is closely connected to the principle of *lärarundantag* (the teacher exception), which grants individual academics professional authority over their teaching practices. While higher education institutions may regulate the scope and extent of teaching duties, decisions concerning *how* teaching is conducted are largely regarded as teachers' responsibilities. Teaching materials, including examinations and assessment tasks, could thus be considered as expressions of academic judgement rather than as merely managerial instruments subject to standardised control. Such a framing positions university teachers as simultaneously course developers, teachers, and examiners, bearing responsibility for decisions concerning *what is taught*, *how teaching is organised*, and *how student learning is assessed*. These responsibilities are explicitly articulated in the Higher Education Ordinance (Högskoleförordningen [Higher Education Ordinance], 1993:100, p. K.6, §18; amended by Förordning 2021:1338),

which assigns teachers the role of examiners holding the authority to assess student performance and determine grades. The Swedish regulatory framework thereby reinforces the central role of teachers in shaping both teaching and assessment practices. Additionally, expectations regarding teaching competence further emphasise the professional responsibility of university teachers. Teachers are expected to exercise independent professional judgement in teaching and assessment and to "independently and jointly with others, plan, implement and evaluate teaching and assessment in higher education with a scientific, scholarly or artistic basis and within their own area of knowledge" (SUHF [Association of Swedish Higher Education Institutions], 2016, p. 2). These expectations underscore teaching as a practice grounded in disciplinary expertise and scholarly standards.

Such professional autonomy is not absolute, as the exercise of professional judgements are expected to take into account the broader social, national and institutional context in which takes place. University teachers are expected to situate their practice within broader societal and institutional objectives by complying with institutional regulations, national quality assurance systems, and public accountability requirements. This dual positioning captures the multiple facets of the university teacher's role in Swedish higher education; as an autonomous academic professional responsible for course design, teaching, and assessment, and as a state-funded employee operating within a regulated public system. These conditions assign university teachers substantial responsibility for the design, implementation, and evaluation of teaching, while simultaneously embedding this responsibility within institutional and national governance structures.

Such specificities of the roles of university teachers in Swedish higher education are central to addressing the aim of my thesis, which foregrounds teachers' perspectives. Examining how teachers perceive and experience AI artefacts in relation to their practices, and how AI mediates their understandings of *what they ought to do* and *how they ought to act*, when balancing sometimes competing demands of autonomy and accountability, provides a distinctive analytical point to approach my research questions. In this sense, Swedish higher education constitutes a particularly relevant and meaningful empirical site for the present research.

# Approaching ethics in higher education and AI in education

*"any discussion with regard to the responsibilities of the university teachers leads us naturally to a discussion of the values that underpin teaching in higher education"*
Macfarlane's (2003, p. 30)

This chapter aims to explore ethical issues related to higher education practice and AI in education (AIED). First, I frame higher education practice as an ethical practice concerned with several core values. Second, I present the current landscape of AIED ethics literature and discuss its gaps in approaching ethics in relation to AI and higher education practice. Finally, I turn to relational ethics as a fruitful approach for understanding ethics in the context of AIED and AI in higher education.

## Ethics in higher education

In this section, I aim to describe higher education practice as an ethical practice concerned with several core values. Framing higher education as an ethical practice relates to a dynamic interplay between institutional norms, disciplinary traditions, and individual and professional virtues. More specifically, this chapter draws primarily on Macfarlane's (2003) *Teaching with Integrity* and Harland and Pickering's (2011) *Values in Higher Education Teaching*, examining how ethics is enacted within university teaching, and how values shape academic practice.

Ethical considerations in higher education are complex, contextual, and deeply intertwined with the values that underpin teaching, learning, and academic life. University teachers navigate moral and ethical dilemmas daily as they balance professional autonomy, institutional expectations, and the diverse needs of students. Harland and Pickering (2011) describe higher education as fundamentally a *values enterprise*. They argue that values form the foundation of all academic activity, influencing both the content and the conduct of teaching and research. According to the authors:

> "Values are an unavoidable result of experiencing university life and therefore a responsibility of the institution and its community. They are the result of an

> education because to educate is necessarily to pass on a set of possibilities and expectations that are selected on value grounds. They are a responsibility because all such possibilities and expectations, and their implicit values base, have some impact on society" (Harland & Pickering, 2011, p. 6).

Harland and Pickering (2011) conceptualise higher education as a moral landscape where values are embedded in everyday actions. Higher education, in their view, is not and cannot be value-neutral, as both teaching and the disciplines in which teaching is situated are value-laden. The authors state that teachers and students continually make value judgments that shape both their academic practice and the broader purposes of higher education, noting that "every choice a teacher or student makes is informed by their values (Harland & Pickering, 2011, p. 3). Values, in this sense, are inseparable from practice. Teaching and learning involve continuous acts of valuing – deciding what counts as knowledge, what is worth knowing, and how that knowledge should be taught and assessed. University teachers, addressing the question 'What's worth knowing?' as part of course development, as well as seemingly technical decisions in course design and assessment, are value-laden.

The same authors identify a set of foundational values for academic practice, including *academic freedom*, *intellectual honesty*, *respect for diversity*, and *critical thinking*. Although they describe such values as constituting a moral framework that guides the university's contribution to society, they caution against viewing such values as universal constants.

> "At one level these ideas might seem reasonable as part of a foundation for the academic profession but closer inspection will inevitably reveal that each will have diverse meanings in different contexts and therefore there is potential for wide interpretation leading to different uses and outcomes. A list such as this lacks a clear referent and purpose, and values need to be understood with respect to the circumstances in which they are realized." (Harland & Pickering, 2011, p. 64)

Thus, values need to be understood and enacted within the situated realities of disciplines, institutions, and cultural contexts. Macfarlane (2003) also presents certain values framed as foundational: *academic integrity*, *tolerance of others* and *mutual respect*. Macfarlane describes academic integrity as "a cornerstone of academic life" (p. 29), essential for maintaining trust and credibility within the academic community and often related to avoiding plagiarism and intellectual theft of writings and ideas. However, as the author states, integrity extends beyond avoiding plagiarism or misconduct; it involves a commitment to truth-seeking and intellectual honesty, demanding "to be assiduous and conscientious in unearthing and using sources of information" (p. 29). Tolerance of others and mutual respect are also described by Macfarlane as central values

of academic life, ensuring or enacting academic freedom, allowing higher education to function as an open space to diverse views and a forum for intellectual debates.

While Harland and Pickering (2011) emphasise the pervasive presence of values, Macfarlane (2003) focuses on how educators engage with ethical dilemmas that arise from those values in practice. He argues that "ethics involves engagement with complex situations and making hard choices" (p. 2), emphasising that university teachers often confront moral tensions in managing their relationships with students and colleagues, in assessment and supervision, and in exercising professional autonomy. Macfarlane (2003) argues that higher education institutions have increasingly responded to (sporadic) ethical challenges by developing codes of conduct and formalised rules. However, he critiques this approach as reductionist, noting that these rules can provide "a benevolent protection from difficult decision making dilemmas" (Macfarlane, 2003, p. 16), but they also risk discouraging the use of personal moral judgment. The overreliance on codified procedures could lead to a culture where ethical complexity is simplified into rule compliance. Hence, ethics in higher education cannot be reduced to compliance with institutional codes or regulations, but rather encompasses broader questions of fairness, integrity, responsibility, and respect within the academic community (Macfarlane, 2003).

At the same time, Macfarlane (2003) recognises that higher education is a "special context" (p. 3), characterised by significant autonomy and privilege. Academics enjoy freedoms in curriculum design, research, and expression, yet this autonomy carries a corresponding obligation for professional integrity and ethical responsibility. This is often more evident in assessment practices, which Macfarlane (2003) identifies as one of the most ethically demanding aspects of teaching. Although assessment is heavily regulated, "the power and individual discretion of lecturers should not be underestimated" (Macfarlane, 2003, p. 71). This underscores the importance of moral judgement and fairness in the everyday work of teaching. Macfarlane's argument culminates in a call for a virtue ethics approach to higher education - one that focuses on the moral character of the teacher rather than adherence to rigid principles. He writes that:

> "rather than a 'rules and regulations' driven approach to ethical issues on campus, what is required is an identification of virtues compatible with reflective professionalism. This requires that the exercise of professional judgement is based on core moral virtues and conceived as a central duty of academic life" (Macfarlane, 2003, p. 127–128).

Ethics in higher education, then, should be seen as a matter of practical wisdom, a form of moral artistry shaped by experience, empathy, and reflection. Overall, Macfarlane (2003) suggests that such teaching virtues align with

Holloway's (1999) view of ethics and morality being as much like an art as like science[10], using the metaphor of jazz music, based on improvisations and responsiveness, but within a certain scale (key, tempo, rhythm) of a common values framework. Such a view in the context of teaching in higher education means:

> "recognizing the importance of basic moral virtues in forming relationships with students and colleagues based on trust and mutual respect. It also means that rational virtues such as a sense of fairness or justice, need to be combined successfully with affective virtues such as sensitivity." (p. 145)

Such an analogy captures the emphasis Macfarlane (2003) puts on university teachers' professional judgment, navigating a dynamic and context-sensitive practice that balances autonomy with accountability, and rules with compassion.

Both Macfarlane's (2003) and Harland and Pickering's (2011) views on ethics in higher education suggest that ethics in university teaching depends on both institutional frameworks and the cultivation of personal virtues. Rules and codes provide necessary boundaries, but it is the educator's character and capacity for reflective judgement that bring those values to life in practice.

Macfarlane (2003) suggests a series of virtues of university teaching illustrated in Table 4.

*Table 4 Macfarlane's virtues for university teachers (Macfarlane, 2003, p. 128)*

| Virtue | Examples of application |
|---|---|
| *respectfulness* | in teaching students and in relations with colleagues |
| *sensitivity* | toward students seeking tutorial support; conducting peer review activities |
| *pride* | in adequate preparation to teach |
| *courage* | to innovate in teaching practice; confront challenging situations with students and colleagues |
| *fairness* | particularly in relation to assessment issues; investigation of complaints about colleagues |
| *openness* | in relation to self, peer and student evaluation of practice |
| *restraint* | in conveying the teacher's ideological and/or theoretical position; checking emotional reactions |
| *collegiality* | in managing courses and invoking consultative processes with students and colleagues |

[10] "Morality is as much an art as a science and it calls for a certain versatility from us, the ability to improvise and respond to actual circumstances and particular situations." (Holloway, 1999, p.16)

Such relational virtues such as *respectfulness, sensitivity, pride, courage, fairness, openness, restraint*, and *collegiality* are several key virtues of teachers underpinning university teachers' practices. These virtues help teachers navigate the emotional and moral complexity of higher education, where, as he notes, "teaching induces a sense of emotional vulnerability" (Macfarlane, 2003, p. 55). Macfarlane (2003) notes that:

> "establishing a teaching relationship based on trust, modelling and encouraging respect for others and caring for students while maintaining a reputation of being fair (or just) to all." (p. 67)

Harland and Pickering (2011) reinforce this relational dimension by framing values as enacted through everyday choices and interactions. "Valuing is about choices. We evaluate as we make decisions about what we do and how we live, and under normal circumstances, our thinking and actions express or reflect our values" (Harland & Pickering, 2011, p. 9). From this perspective, ethics is not merely a matter of compliance but of conscious valuing as an ongoing reflective practice that informs teaching, research, and collegial relationships.

Finally, universities tend to increasingly require university teachers to draft their philosophy of teaching statements as part of recruitment, promotion, or evaluation processes. These statements serve as a form of professional reflection that articulates a teacher's core beliefs, values, and goals about teaching and learning (Caukin & Brinthaupt, 2017). A philosophy of teaching statement, according to Caukin and Brinthaupt (2017), offers insight into how teachers view the purpose of education and their role in the classroom, and encourages reflection on the ethical commitments and pedagogical choices. This process aligns closely with Macfarlane's notion of reflective professionalism, where teachers make morally informed judgments grounded in their experience and values. Similarly, it echoes Harland and Pickering's view of higher education as a "values enterprise" in which every aspect of teaching is value-laden. Through writing and revising their philosophy of teaching statements, university teachers engage in a form of ethical self-examination, clarifying how personal beliefs align with institutional values and broader societal purposes of higher education.

Emerging technologies mediating higher education practices actively participate in, reshape, and potentially disrupt teachers' ethical and valuing landscapes. These technologies amplify – or, in some cases, obscure – ethical concerns, challenges, and dilemmas that are not entirely new to higher education, but are instead deeply rooted in its practices. Therefore, ethical issues arising within AI-mediated practices need to be approached not as unprecedented disruptions, but as continuations and reconfigurations of long-standing ethics in university teaching, refracted through new technological mediations.

## AI in education ethics

In this section, I briefly present the current landscape of AI in education (AIED) ethics literature and discuss its gaps in approaching ethics in relation to AI and higher education practices. The main aim of this section is to offer an overview of how AIED ethics is currently framed across the broader literature, including academic research and policy documents produced by international, national, public and private stakeholders. At the same time, it highlights the limitations of these approaches in allowing for a comprehensive understanding of ethics in relation to deploying AI artefacts in higher education. Rather than attempting to identify or consolidate a set of 'key' AIED ethical issues, the focus of this section lies on the gaps that such issue-based and classificatory approaches tend to leave open.

AI technologies are often framed and understood as pragmatically capable of functioning in social settings through quantifiable calculative processes and problem-solving logics, given certain inputs, data and instructions (Berendt, 2019; Selwyn, 2022). Such logics, of seeing AI artefacts being designed and deployed within the frame of close-ended tasks, along with the broader datafication (and *AI-fication*[11]) of society, have led to the a solution-oriented and challenge-driven approach to exploring and understanding ethics in the context of AI and AIED. This is evident by the plethora of ethical codes, guidelines and frameworks that have been developed (Jobin et al., 2019) from academia (Slade & Prinsloo, 2013), national bodies (Bakewell et al., 2018), industry (Microsoft, 2023), and international organisations (European Commission, 2022; UNESCO, 2022). However, these documents remain voluntary, lacking enforcement, and often conceptualise AI as a technical system disconnected from sociocultural contexts (Hagendorff, 2020; McNamara et al., 2018). Whittlestone et al. (2019) further criticise them for vague definitions, neglect of value tensions, and limited grounding in evidence. As a result, existing frameworks provide only high-level orientation and fail to capture the contextual dilemmas of specific AIED artefacts such as AGS and GAI chatbots (Floridi, 2019). Generic recommendations cannot accommodate the diversity of legal, cultural, and socio-economic conditions shaping educational practices (Slade & Prinsloo, 2013).

Such common framings of AI and AIED ethics in the literature – often presented as lists of ethical issues framed as challenges to be addressed through technical solutions – tend to position ethics as a matter of applied ethics. In this view, ethical engagement is reduced to posing ethical questions and calling for moral and normative judgments in relation to the development and deployment of AI in society, including education. While understanding AI and

[11] I use the term *AI-fication* in this thesis as refering to a process in which social practices, institutional arrangements, and forms of governance become increasingly organised around the presumed capacities of AI technologies.

AIED ethics through an applied ethics approach may be convenient and effective for addressing problems that arise in practice, as Selwyn (2022) cautions, "the moral conundrums and challenges that underpin AI ethics in education are obviously not wholly reducible to sets of discrete procedural challenges" (p. xiv). Ethical discussions in AIED and its related field of learning analytics have therefore concentrated on institutional data governance (Cerratto Pargman et al., 2023; Cerratto Pargman & McGrath, 2021), while moral and pedagogical questions in teaching and learning remain underexplored (Cerratto Pargman & McGrath, 2021; Pardo & Siemens, 2014; Schiff, 2022; Slade & Prinsloo, 2013). This leaves a significant gap in the AI and AIED ethics literature in terms of critically understanding and engaging with ethics beyond procedural concerns.

Holmes and Porayska-Pomsta (2022) note that, while AIED research and development have not necessarily been unethical, the field lacks systematic engagement with ethics, resulting in fundamental ethical questions about the use of AI in education remaining underexplored and only partially articulated. More specifically, they state that:

> "[L]ittle exploration of whether, how, or why technology – and especially AI technologies – may fundamentally change the way that people think and behave (broadly speaking, human psychology) has been done within the AIED context beyond evaluations of the effects of the specific pedagogical designs of the AIED systems they use on students' domain-specific problem-solving skills." (Holmes & Porayska-Pomsta, 2022, p. 12)

According to the authors, little attention has been given to how technologies affect the existential space and values when engaging with AIED, leaving underexplored questions regarding the values and visions encoded into AI artefacts in education. In this light, Kizilcec and Lee (2022) discuss *Algorithmic fairness* and the ethical footprint of AI systems in education in connection to society, arguing that "the increasing use of algorithmic systems in education raises questions about its impact on students, instructors, institutions and society as a whole" (p. 175). Kizilcec and Lee (2022) encourage academic policymakers to engage with a critical analysis of AIED, posing crucial questions about the meaning of equity in AI systems, such as: How much and under what circumstances do such technologies benefit education stakeholders? What characteristics of algorithmic systems in education are associated with greater benefits? What counts as a beneficial impact on education stakeholders? What is the AI system's definition of student success/failure? How were students in the training dataset selected? (p. 195, 205). Considering such questions and critically examining the uses and misuses of AI in education is necessary, as these issues are inherently ethical, political, and entangled with broader practices and structures of power (Verdegem, 2021).

Furthermore, a significant gap remains in the literature, as little empirical attention has been given to ethics. In particular, empirical research on AIED has tended not to critically examine the ethical implications, challenges, and risks associated with the introduction of AI into higher education institutions (Bond et al., 2024; Zawacki-Richter et al., 2019). In their review, Zawacki-Richter et al. (2019) found that only two out of 146 articles on AI in higher education address ethics, focusing merely on privacy concerns and the costs (financial) of AI development that not many public higher education institutions could potentially afford. Reflecting on their findings, Zawacki-Richter et al. (2019) discuss several potential ethical implications and risks when integrating AI in higher education institutions, focusing on replacing teaching with automated AI-powered solutions in case of institutions' financial constraints, or the 'fear' teaching and administrative staff might face by being replaced by AI systems. Additionally, the same authors discuss the potential issues related to privacy and data protection since AI systems (so far) seem to require huge amounts of data (Zawacki-Richter et al., 2019). Similarly, Cerratto Pargman and McGrath (2021), in their review on the ethics of Learning Analytics (LA)[12] in higher education institutions, found that the work to date is of a predominantly conceptual character, and identified that transparency, privacy, and informed consent are the most frequently addressed ethical issues, while research on justice, equality, bias, ethical dissonance, moral discomfort, and intellectual freedom is scarce (Cerratto Pargman & McGrath 2021). Furthermore, Cerratto Pargman and McGrath (2021) discuss the multiple layers of such ethical issues as the purposes driving the use of the LA systems in education, the power entrenched in who gets to collect, store, and control institutional and student data, and the cultural, sociotechnical, and pedagogic contexts involved.

This section concludes by highlighting several persistent gaps in the AIED ethics literature. First, dominant framings of AIED ethics remain largely issue-based, procedural, and solution-oriented, often approaching ethics as a form of applied ethics concerned primarily with governance, compliance, and technical risk mitigation. Second, ethical discussions in AIED have focused predominantly on data governance, privacy, and transparency, leaving questions of values, power, equity, pedagogy, and human experience comparatively undertheorised. Finally, despite growing scholarly interest in AI in

[12] Learning analytics (LA) and artificial intelligence in education (AIED) represent distinct but overlapping fields. Both rely on data-driven methods to improve learning and teaching, but they differ in focus and scope. AIED typically involves the development of systems that interact with learners or teachers (Luckin et al., 2016), while LA emphasises the collection, measurement, and interpretation of learner data to understand and optimise learning environments (Siemens & Baker, 2012). In practice, the two fields increasingly intersect, as advances in AI techniques enhance data analysis in LA, and LA insights inform the design and evaluation of AIED systems.

higher education, there remains a notable lack of empirical research that critically examines issues of ethics in the context of AI deployment in educational practice. On this basis, these gaps point to the need for more context-sensitive, empirically grounded, and critically informed approaches to AIED ethics that move beyond classificatory frameworks and procedural checklists toward a deeper engagement with ethics in relation to higher education practices.

## Relational ethics

This section turns to relational ethics as an alternative approach in exploring ethical issues in relation to AI as they emerge within higher education practices.

Relational ethics offers an alternative to dominant procedural and principle-based approaches to ethics by shifting attention from abstract norms and compliance-oriented frameworks to relationships, context, and practice as it unfolds in situated settings (Bergum & Dossetor, 2005; Birhane, 2021). Rather than conceptualising ethical issues as problems that can be resolved through the application of predefined principles to discrete technologies, relational ethics encompasses ethical and moral concerns as emerging through continuous interactions among human and non-human actors. From this perspective, ethical questions cannot be meaningfully separated from the social, institutional, historical, and material conditions in which technologies are designed, developed, and deployed. Along these lines, Mittelstadt (2019) states that ethics refer to complex and difficult debates that are not "meant to be easy or formulaic" (p. 505). More specifically, the author concludes that:

> "Intractable principled disagreements should be expected and welcomed, as they reflect both serious ethical consideration and diversity of thought. They do not represent failure, and do not need to be 'solved'. Ethics is a process, not a destination." (Mittelstadt, 2019, p. 505)

Such a view also aligns with Slattery and Rapp's (2002) discussion of ethics as emerging through multi-agentic relations, actions, and reactions, arguing that:

> "[E]thics can be understood only in the context of the messy and ambiguous struggle to live meaningful and consequential lives in the midst of human aspirations, desires, frailties, and shortcomings. … Ethics involves every dimension of living and being in the world." (p. 21)

In this view, AI and AIED ethics are situated in practices, contexts, and stakeholders. Exploring ethics in these terms therefore requires considering the messy reality of ethical debates, values negotiations, and moral judgements

that are relevant to such practices, contexts, and stakeholders. Ethics is thus understood as dynamic and evolving, shaped by relationships, purposes, and values that are enacted in specific contexts rather than fixed in advance.

Central to a relational ethics approach is the recognition that actors are not isolated entities but are mutually constituted through their relationships, and that ethical responsibilities arise from these interdependencies (Decuypere & Simons, 2016; Slattery & Rapp, 2002). In educational contexts, this entails examining how AI technologies reconfigure relationships between students, teachers, institutions, and data infrastructures, and how these reconfigurations shape agency, power, trust, and professional judgement. Relational ethics therefore directs attention to questions of how ethical issues are enacted in practice, how values are amplified or diminished through sociotechnical arrangements, and how moral responsibilities are negotiated across time and place. By foregrounding context, interdependency, uncertainty, and vulnerability (Bergum & Dossetor, 2005; Pollard, 2015), relational ethics enables a more nuanced and practice-oriented engagement with ethics in AI-mediated educational practices than approaches centred primarily on compliance, checklists, or technical risk mitigation (Birhane, 2021).

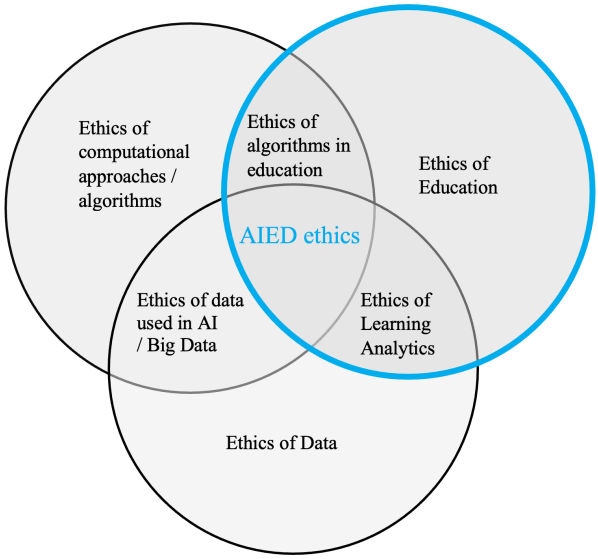
On this note, Mittelstadt describes AI ethics as "a microcosm of the political and ethical challenges faced in society" (2019, p. 505). Echoing this view, Selwyn (2022) argues that AIED ethics is indistinguishable from education ethics. More specifically, the author states that:

> "talking about 'AI ethics in education' we are primarily talking about educational ethics and societal ethics -issues and debates that have engaged diverse groups and communities for centuries." (p. xvi)

According to this view, the ethical issues that emerge when AI is introduced into higher education practices are not novel, but relate to long-standing ethical dimensions of such practices.

In line with this approach, Holmes et al.'s (2022) work on AIED ethics foregrounds education ethics and conceptualises the ethical considerations of AI systems regarding data, algorithms and education practices. Arguing that exploring AIED ethics requires a multidisciplinary approach, Holmes et al. (2022) present a draft but rather comprehensive framework involving an amalgamation of different ethical concerns in relation to data, algorithms and education (see Figure 1).

*Figure 1 AIED ethics – modified and adapted from Holmes et al. (2022, p. 521)*




In this framework, Holmes et al. (2022) highlight the need to account explicitly for the ethics of education, which have often been overlooked in earlier AIED ethics research. More specifically, the following areas are distinguished by Holmes et al. (2022), namely: *ethics of data,* which speaks to how data is collected and managed, *ethics of algorithms,* which addresses how decision-making models are developed and tested, and *ethics of education,* which refers to the implications such systems may have on professional practice. Holmes et al. (2022) discuss ethics of data, referring to ethical concerns related to the collection, access and storage of data, data ownership, privacy, accountability and responsibility. Ethics of algorithms and computational approaches in education are connected to the choices made for decision-making models' development. More specifically, Holmes et al. (2022) pose questions regarding issues related to the analysis and interpretation of data, conscious and unconscious biases and potential harms to civil rights and students' discrimination. Finally, the ethics of education regards questions related to the purpose of learning, the choice of pedagogy, the role of technology with respect to teachers and access to education (Holmes et al., 2022). Holmes et al.'s (2022) approach to AIED ethics aligns closely with relational ethics, as it emphasises the importance of understanding ethics as situated in contexts, practices, and relationships, and argues for a shift in focus toward the ethics of educational practices themselves, rather than treating ethics solely as a matter of data governance or algorithmic design.

In this thesis, I place emphasis on how ethical issues emerge when AI artefacts are introduced into higher education practices, using a relational approach to ethics informed by relational ethics theory. Building on Birhane's

(2021) argument that dominant rationalist approaches to ethics are *acontextual* and offer technical fixes to social problems, I draw on a relational ethics approach that situates ethics in lived practices, oriented toward justice and tangible improvements in communities. A relational approach to ethics situates technologies within the broader social and educational contexts in which they are used. It focuses on how technologies mediate relationships and shape practices, while foregrounding the values, rights, and responsibilities that arise from these interactions. Rather than asking only 'what' or 'who' is implicated, relational ethics also focus 'how' and 'where' ethical issues emerge, remaining attentive to the specificities of each setting (Decuypere & Simons, 2016). Additionally, quoting Verbeek (2005, p. 117), "technologies cannot be separated from their use contexts… they are what they are only in their use." This view underscores that ethical concerns cannot be abstracted from the social practices in which technologies are embedded. Relational approaches, therefore, provide a dynamic and context-sensitive lens, one that treats ethics as situated, evolving, and interconnected, and capable of revealing aspects often overlooked by universal or abstract frameworks.

# Theoretical underpinnings

*"technology forms one of the tissues of meaning within which our existence takes shape"*
Verbeek (2011, p. 155)

This chapter presents the theoretical and conceptual foundations of my thesis by offering an overview of postphenomenology and technology mediation theory. Postphenomenology provides a framework for approaching the research questions, situating AI technologies not as neutral tools but as 'mediators' of human experience and lifeworld relations. This perspective foregrounds the co-constitutive role of technology in shaping perceptions, actions, and meanings. Building on this foundation, technology mediation theory extends the analysis into the domains of practice, values, and ethics, providing a conceptual and analytical orientation for examining how technologies are perceived, participate in and transform human–world relations.

## Postphenomenology

My research draws on postphenomenology, which provides a framework for understanding the complex and constantly evolving experiences, interactions and relations between humans, technology and the world.

Postphenomenology is a philosophical perspective that has significantly influenced the philosophy of technology during the past three decades, focusing on studying how humans relate to technology (Rosenberger & Verbeek, 2015). This philosophical perspective prompts researchers to go beyond views of technology perceiving it as mere tools, acknowledge its profound influence on existence and experience, and focus on understanding how technologies shape and are shaped by our choices, actions, and our perception of the world. Postphenomenologists also inquire into how technologies configure and shape our purposes and, in doing so, exerting influence on us, like for instance on our perspectives on politics, ethics and aspects of everyday life.

Postphenomenology develops out of the phenomenological tradition but reconfigures its central task. While phenomenology seeks to describe human

experience and the disclosure of the world, postphenomenology shifts the emphasis toward the relational character of human–world experience as it is mediated by technologies. As Verbeek (2005) explains, this approach draws simultaneously on phenomenology's descriptive orientation and pragmatism's empirical concerns, and takes *mediation* as its fundamental concept for understanding human-world relations. 'Post'- in postphenomenology indicates its distance from classical phenomenology's romanticism of technology as well as its relation with postmodernism and poststructuralism in terms of studying the relation between the subject and the object. Drawing on phenomenology, Ihde (1990, 1993), who can be considered the founder of this research tradition, has developed a framework of concepts that describe human experiences in the lifeworld through their relation with technology, where technology is seen as "shap(-ing) our relations to the world, rather than merely distancing us from it" (Rosenberger & Verbeek, 2015, p. 11). Ihde (1990) aims to describe such experiences through his analysis, by going beyond classical phenomenological accounts that tended to discuss technology in abstract or romantic terms, or as a threat to human authenticity, and directs attention to concrete artefacts and practices, showing that technologies are constitutive of how humans perceive, experience and make sense of the world and their practices.

Additionally, postphenomenology intensifies one of phenomenology's core insights: the intentional relation between subject and object, which cannot be separated and is rather co-constitutive. According to postphenomenology, human-world relations should not be understood as interactions between already-formed subjects and a pre-given world of objects; rather, they are the sites where both the 'objectivity' of the world and the 'subjectivity' of those experiencing it are constituted (Verbeek, 2011, p. 16).

> "Phenomenology thus overcomes the dichotomy between subject and object, humans and world, by replacing it with a mutual interrelation. Human beings are unthinkable apart from a relation to the world, which they continually experience and in which they realize their own existence. […] a new interpretation of phenomenology can take this a step further by emphasizing that subject and object constitute each other. Not only are they intertwined, but they coshape one another. Human beings can only experience reality by relating to it, which does not involve any reality-in-itself but rather reality-for-them." (Verbeek, 2005, p. 110, 112)

In this respect, postphenomenological investigations of human-world relations mediated by technology, which, to Ihde, constitute several schemes of human-technology-world relations, together with focusing on the technologies (or things) themselves, can provide several analytical insights into human subjectivities. Furthermore, motivated by postmodernism and by a desire to overcome the sharp subject–object split, postphenomenology is characterised by a relational ontology, claiming that "reality arises in relations, as do the

human beings who encounter it" (Verbeek, 2005, p. 113). In such a relational ontology, reality, as it matters for human beings, arises in relations – in the concrete, mediated entanglements between humans, artefacts, and world – and, likewise, human subjects take shape in and through those same relations; in other words, there is no access to a reality "in itself" that would be prior to mediation, but neither is there a sovereign subject inventing reality out of whole cloth. As technological mediations participate in shaping both subjectivities and the realities they encounter, postphenomenology becomes directly relevant to studying the ethical issues in relation to technological artefacts, offering a way to analyse how technologies carry moral significance by actively configuring how people experience, interpret, and act in the world (Verbeek, 2011).

Additionally, Sergeeva (2023) points out that instead of assuming technology's universal power separating humanity from its natural, authentic lifeworld, postphenomenology focuses on investigating 'technology-in-the-particular' – the specific uses of particular technologies in situated practices, aiming to explore how technologies play a constitutive role in human existence through the lens of mediation. Postphenomenological research involves examining how situated experiences differ under conditions of technological presence or absence, specifically focusing on transformations of perception, brought about by mediation.

Postphenomenology offers a particularly suitable conceptual orientation for studying technologies in education insofar as it foregrounds technological mediation as a constitutive dimension of educational practice, rather than treating technologies as neutral tools or external variables. According to Aagaard (2017), educational research has traditionally been sociocentric, often "blackboxing" the role of material artefacts despite their pervasive presence in classrooms and academic work. Postphenomenology responds to this gap by focusing on how technologies mediate human perception and action, thereby shaping how educational practices are experienced and enacted. By emphasising multistability, embodiment in a broader sense, and the non-neutrality of artefacts, postphenomenology enables researchers to explore how technologies shape possibilities for teaching and learning without reducing these effects to technological determinism or individual intentionality. In this way, postphenomenology provides a nuanced framework for examining human-technology relations *from within* educational practice, while remaining attentive to context, use, and lived experience.

Although postphenomenology does not offer strict guidelines for researchers on how to design their study, there are certain 'rules' that previous postphenomenological studies have followed (i.e. in Ihde, 1993; Rosenberger, 2012; Selinger, 2006):

a) They examine technology in terms of relations between human beings and technological artefacts, focusing on how technologies mediate human-world relations (relational ontology); and
b) Combine philosophical analysis with empirical investigations, taking as a starting point a specific technology in place of a situation (philosophy *from* technology) (Rosenberger & Verbeek, 2015, p. 9-10).

With regards to the methodological characteristics of postphenomenological research, Rosenberger and Verbeek (2015) argue that it is highly diverse and context-sensitive and that it could be seen as "empirical philosophy" since its conceptual analysis can only occur within a context of actual technological practices and artefacts, investigating the implications of such practices and artefacts (p. 30). For instance, studying Learning Management Systems (LMS) with a postphenomenological inquiry could focus on how these systems help to shape new forms of interactions between teachers and students and create room for new ethical questions on how LMS-on-student-progress-informed teachers should act. In an attempt to list the characteristic elements that an empirical-philosophical postphenomenological approach includes methodologically, Rosenberger and Verbeek (2015) note that postphenomenological studies:

- "focus on understanding the roles and characteristics that technologies play in the relations between humans and world, and on analyzing the implications of these roles" (p. 31);
- "include empirical work as a basis for philosophical reflection (…) to investigate the character of the various dimensions of the relations between humans and these technologies, and their impact on human practices and experiences" (p. 31);
- "investigate how, in the relations that arise around a technology, a specific "world" is constituted, as well as a specific "subject" (p. 31); and
- "make a conceptual analysis of the implications of technologies for one or more specific dimensions of human-world relations- which can be epistemological, political, aesthetic, ethical, metaphysical et cetera" (p. 31).

Aagaard (2017) outlines two methodological approaches for conducting postphenomenological research. The first involves an in-depth examination of the typical use of a specific technology, while the second entails a critical comparison of multiple versions of a technology as they are enacted in practice. Both approaches resonate with Verbeek's (2011) understanding of human freedom in relation to technological artefacts. As Aagaard (2017) explains:

> "Technologies do not determine what we do: They set up spaces of possibilities that enable and constrain certain actions and perceptions, but the room for maneuver within these spaces is what constitutes human freedom. Human freedom is not an absence of technological influences, but the practice of coping with such influences." (p. 527).

With regard to the first methodological approach, Aagaard draws inspiration from Adams and Thompson's (2011) notion of *interviewing objects*, exemplified in their study of how PowerPoint mediates teachers' modes of presenting knowledge. This approach directs the researcher's attention to the affordances of technologies and to the experiential spaces they shape within professional practice. The second approach, by contrast, focuses on how technological artefacts exhibit multiple stabilities depending on how they are interpreted and enacted by different users across diverse contexts.

My research aligns partially with the first methodological approach outlined by Aagaard (2017) and with Adams and Thompson's (2011) fourth heuristic of "recognizing the amplification/reduction structure of human–technology relations" (p. 741). Methodologically, this means that I do not treat automated grading systems (AGS) and generative AI (GAI) chatbots as neutral tools, nor as determining forces, but as mediating technologies that simultaneously foreground certain pedagogical possibilities while backgrounding others. Accordingly, my studies examine how these technologies are perceived and experienced by university teachers within the context of their practices. Rather than systematically mapping amplification and reduction effects, the heuristic functions as a sensitising lens that orients attention to how technological mediations open up certain possibilities while constraining others. This perspective supports an exploration of a range of issues, including ethical, pedagogical, and practical concerns, that emerge as teachers engage with AGS and GAI chatbots in higher education.

## Technology mediation theory

Central to postphenomenology is the concept of *mediation*, which allows scholars to move beyond two dominant yet ultimately reductive approaches to technology: the *instrumentalist view,* which conceives of technological artefacts as neutral tools that merely serve human good or bad intentions, and the *substantivist view*, which perceives technology not as neutral but as an autonomous and determining force that shapes society and culture from the outside. Rather than aligning with either of these perspectives, technology mediation theory highlights the ways in which technologies actively participate in shaping human relations – with the world, with one another, and with technologies

themselves – so that experience is never simply direct but always technologically mediated in some form (Verbeek, 2005). At the same time, the mediating role of technology is not fixed or predetermined, for technologies do not exist in isolation but only within specific contexts of use and interpretation, which means that their significance emerges in the interplay between artefact, practice, and meaning (Verbeek, 2005). In this sense, technology mediation theory not only resists simplistic classifications of technology as neutral or determining but also foregrounds the dynamic and situated character of technological influence[13]. Verbeek (2005), unfolding the complexity of the concept of mediation and its distinction with transformation as a process of change or influence between fixed entities (subject and object), states that:

> "the relation between subject and object always already precedes the subject and the object themselves, which implies that the subject and the object are mutually constituted in their interrelation. In any relation between subject and object, both are brought into existence in a specific way, and both subjectivity and objectivity acquire specific shape. When analyzing the mediating role of artifacts, therefore, this mediation cannot be regarded as a mediation "between" subject and object. Mediation consists in a mutual constitution of subject and object."(pp. 129–130)

Ihde (1990) advances his account of technology's place in human–world relations by asking how those relations are configured once technological artefacts are involved and differentiates three forms of humans relating with technology. The first form regards *relations of mediations* where perception itself is routed through a device, so that access to the world is achieved by way of an artefact-wearing corrective lenses and using a thermometer to read the temperature are common examples of this perceptual mediation. Second, *alterity relation*s regard the relations not with the world through technology but with the artefact itself – interacting with a system exemplifies this orientation toward the artefact itself. Third, *background relations* name the relations to reality where technologies operate in the background without becoming focal, such as the refrigerator keeping cool the stored food. Using the representation of

[13] "That "the things themselves" are accessible only in mediated ways does not interfere with our ability to say something about the roles that they play, thanks to their mediated identities, in their environment. And it is precisely the postphenomenological perspective that offers a new way of so doing." (Verbeek, 2005, p. 113) According to Verbeek, although post-phenomenology claims that technologies as 'things themselves' exist only within mediations, can still be considered as adequate lens to study the *influence* or the *roles* of technologies in practices.

Human — Technology — World

as a starting point for describing the technological mediation of human experiences in the lifeworld, Ihde (1990) uses '—' to describe relativity, '→' to describe intentionality and parentheses to describe conjunctions.

Starting with relations of mediations, the ***embodiment relation*** is expressed as:

*(Human — Technology) → World*

Ihde (1990) describes how humans experience the world *through* a technology that shapes their perception. Using a pair of eyeglasses as an analytical device, Ihde explains that a person wearing eyeglasses experiences seeing the world through the transformative mediation of eyeglasses' lenses that allow, augment or support human seeing. Another example related to the scope of this thesis is how AGS mediate assessment when used by teachers, where the teacher grades students' work through an AGS, which could be represented as (teacher — AGS) → students.

Second, the ***hermeneutic relation***, represented as:

*Human → (Technology — World)*

describes how humans experience a transformed encounter with the world by reading and interpreting technology. One example of such a relation is when using a thermometer to take the temperature. The thermometer, in such a case, presents a readable-to-humans way of the world in terms of coldness and warmth. Using an example of hermeneutic relation in the context of AI and higher education, we can think of how a student is informed about the average number of students passing a course in an LMS, where the student reads this information mediated by the LMS interface. This example could be represented as student→ (LMS — course information).

Then, the ***alterity relation*** is represented as:

*Human → Technology (— World)*

describes how humans relate to technologies themselves, often functioning for a specific purpose, representing an automation, and being somehow similar to an interaction (or often transaction) between humans. A lot of examples of such relations can be found in our current societies where financial transactions are handled via ATMs, GPS roadmap guidance, automatic answering home telephone machines and online customer service run by chatbots. In this type of relation, Ihde (1990) explains the technological mediation of humans with (an often intended anthropomorphic) 'otherness'. In the context of this thesis, a relevant example is the relation between a student and a chatbot provided by a higher education institution to answer questions about course literature. In this case, the student relates to the technology as a quasi-other that mediates access to the course literature. This example could be represented as student→ GAI chatbot (— course literature).

Finally, the ***Background relation***, represented as:

*Human — (Technology/World)*

describes how technologies existing in the surroundings mediate experience by shaping human's environmental context. Examples of such a relation can be found in how humans tend to use refrigerators, automatic lighting and thermostats in heating devices. In such cases, humans experience a technology-controlled environment in which they usually intentionally rely on and wish not to observe. Such an environment is observable or reminded when the absence or withdrawal of the technological mediation. One example of such relation connected with AI and higher education can be found in how teachers (and students) perceive wireless internet provided in the higher education institution's spaces which is usually taken for granted, is widely used during classroom activities and is only observable when the connection speed is slow or the WiFi system is completely off. This example could be represented as teachers — (wireless internet/classroom activity)

Building on Ihde's relations, Verbeek (2008) also describes the cyborg or hybrid human-technology relations, referring to ***fusion relation*** which is similar to embodied relation but actually merge with humans, can be represented as *(Human/ Technology) → World* and examples of such relations can be found in the use of cochlear implants by deaf humans that mediate hearing sounds. Another, type of relation described by Verbeek (2008) is the ***immersion relation*** where technologies merge with the environment and humans experience the world mediated by a series of bi-directional interactions with technology, which can be represented as *Human ↔ Technology/World* and smart devices and IoT home devices interact with humans on specific purposes. Lastly, Verbeek (2008) describes the ***augmentation relation*** which

adds a 'second layer' to the human experience in the world using the example of Google glasses, where a certain technology shapes an alternative augmented perception of a selected view of the lifeworld.

In this thesis, such types of human–technology relations described within technology mediation theory are used as an analytical orientation rather than as a classificatory scheme. While the existence of different forms of mediation is taken into account in conceptualising the role of technology, I do not aim to map, categorise, or exhaustively identify the specific ways in which teachers relate to AGS or GAI chatbots. Additionally, in this thesis, mediation is not understood solely as arising from the concrete use of a technology in a specific activity. Rather, drawing on technology mediation theory, I attend to how the mere presence and perceived availability of AGS and GAI chatbots can mediate educational practices by unsettling established ways that teachers relate with their roles, practices and students.

## Multistability & Transparency

Postphenomenological research on technological mediation foregrounds several key terms related to the materiality of technological features mediating experience, including multistability and transparency.

> "There is, however, one pitfall that needs to be avoided in this analysis of the ability of artifacts to coshape the relation between human beings and world: this ability must not be conceived as an intrinsic property of the artifact itself." (Verbeek, 2005, p. 117)

Artefacts' multistability recognises the constraining nature of technology alongside its openness, allowing for multiple uses, implying that the 'same' technology can exhibit various 'stabilities' when examined within the practices of its use in specific historical settings (Sergeeva, 2023). These stabilities may refer to the multiple functions, uses and roles of a certain technology dependent on and defined by the context in place. In this thesis, examining AGS and GAI chatbots' mediations of teachers' practices, I explore how such stabilities of these artefacts emerge, in terms of their experienced by the teachers functions, uses and roles.

Additionally, Ihde (1990), discussing the forms of human/technology–world relations, problematises transparency in terms of the embodied and/or background technologies. More specifically, transparency refers to a quality or characteristic of technology and its relationship with human users. It represents a state where the technology becomes so seamlessly integrated into the user's experience that it effectively disappears from their conscious awareness. When technology achieves transparency, it no longer feels like a separate entity or tool but instead becomes an extension of the user's senses and abilities, or like a natural part of their body selves. In the

context of this thesis, transparency is treated as relatively high, insofar as the study focuses on technologies in their emergence and attends to how teachers experience such an emergence and perceive the role of the technology in relation to their practices.

## Technopractices & Technoperformances

What matters for postphenomenology is not an abstract account of technology as such but attention to specific technologies in use. Human–world relations are enacted in practice, and it is in concrete, material interactions that mediation can be studied. In this sense, postphenomenology is a practical philosophy: it examines how particular technologies, in situated practices, mediate experience, configure subjectivities, and contribute to the emergence of meaningful worlds (Rosenberger & Verbeek, 2015).

> "It is in practices of interacting with technologies where the phenomenon of technological mediation occurs and can be studied. Human-world relations are practically "enacted" via technologies. (…) The materiality of technology can be studied best in concrete, practical situations of use. This is beautifully reflected in the Greek word "pragmata" which means "things", but which is closely connected to the word "praxis", which means "practice". Postphenomenology is the practical study of the relations between humans and technologies, from which human subjectivities emerge, as well as meaningful worlds." (Rosenberger & Verbeek, 2015, p. 12)

Postphenomenological research in education emphasises embodied and situated experiences, attends to concrete practices, and remains grounded in empirical cases. Studies that investigate technological mediations in classrooms focus not on abstract "technology", but on how concrete artefacts shape lived educational practices. For example, Aagaard's (2015) study discusses technology mediated impatience, showing how students' use of personal devices can mediate attention and reconfigure classroom engagement. Similarly, Voordijk and Vahdatikhaki (2022) demonstrate how virtual reality technologies mediate the relation between engineering trainees and construction practice, offering new ways of experiencing and enacting professional learning. These cases highlight that technological mediation is not peripheral to educational practice but constitutive of how teachers and students relate, act, and make sense of their activities.

Furthermore, Coeckelbergh (2019), examining human-technology relations, explores technoperformances as an attempt to "understand and evaluate what we do with technology and what technology does with us" (Coeckelbergh, 2019, p. 1). Furthermore, Coeckelbergh's study on technoperformances describes the ways in which technology contributes to meaning-makings and highlights the transformative potential of technology

to shape human experiences. Understanding technology use as a performative process, wherein both humans and technologies actively participate, allows for conceptualising the interactions between humans and technologies as dynamic, embodied, moving, social, improvisational, and temporal. Although Coeckelbergh (2019) introduces the concept of technoperformance as a metaphor to analyse the use of technology in performance arts (dance, drama, music), more examples can be found in education and everyday life, such as the use of Grammarly, providing grammatical and syntactical assistance when writing texts and auto-correct function, predicting and suggesting words when drafting messages. These examples of technoperformances showcase that humans and technology perform together in a hybrid way. One interesting aspect of technoperformance regards the outcome of such human/technology acts, which makes it impossible to determine and distinguish human and technology contributions to the final output.

Finally, as Coeckelbergh (2020) points out in technoperformances, "technology is not merely a tool but also takes on a stronger, often non-intended role: not so much as 'mediator' but as choreographer, director, and conductor of what we experience and do" (p. 1). In such cases, when we cannot distinguish whether an idea, a decision or an output was made by a human or a machine, several questions arise such as 'who holds the ownership of the output?' and 'who has the responsibility for the consequences?'.

The relevance of these perspectives becomes particularly apparent in relation to AGS and GAI chatbots in higher education. These technologies are not approached as abstract systems, but as concrete artefacts and in relation to teachers' practices, where they become involved in the ways teaching, assessment, and feedback are carried out. From a postphenomenological perspective, teachers' engagements with AGS and GAI chatbots can be understood as technopractices and technoperformances, in which human and technological contributions are closely intertwined. In such situations, outputs such as feedback, grades, or learning materials often emerge through hybrid human–technology performances, making it difficult to clearly distinguish between human and technological contributions.

## Technomorality

Verbeek (2005) in *What Things Do,* establishes a firm discussion on moral machines and how technologies and human social practices are co-constitutive, arguing that not only humans' values are reflected in the design of technologies but also the technologies themselves determine actions and affect humans' perspectives of reality. Social practices, including moral practices, have been always inextricably linked to technology (Vallor, 2016). Social, political, economic, and educational histories have been

shaped by technological practices, from human earlier ancestors' uses of stones to light fire, sharpen weapons, build houses and draw scenes from their everyday lives, to today's digitised professional, educational, financial, and social realities. As Vallor (2016) also argues:

> "(e)thics and technology are connected because technologies invite or afford specific patterns of thought, behavior, and valuing; they open up new possibilities for human action and foreclose or obscure others" (p. 3).

Examples of such possibilities of allowing, shaping, framing and limiting human actions can be profound with forms of communication between humans involving technologies. These new possibilities that technologies open up for humanity include calls for moral actions and set the ground for new perceptions of realities.

Vallor (2016) on *technomorality* and *technomoral choices,* discusses that human ethical decisions have become increasingly intertwined with technology. Decisions regarding how to live a moral life have been profoundly shaped by technological developments and the affordances that these technologies provide. These affordances refer to the possibilities and capabilities made available by technology shaping the landscape of options and behaviors, influencing how humans engage with the world and with one another. One example related to social media regards the ethical implications and the new ethical questions that arise when participating in such digital social space. Is it ethical to share someone else's photo without their consent? Is it ethical to depict someone's reaction by taking a screenshot? Such questions and concerns were less prominent prior to social media, when human interactions primarly took place in physical spaces. Decisions about privacy, consent and information sharing can be described as technomoral decisions, according to Vallor (2016). From this perspective, human judgements about what is ethically acceptable or wrong are not separated from the technologies that are in use, as these technologies play a significant role on ethical decision-making.

Verbeek (2011) on *Moralizing Technology: Understanding and Designing the Morality of Things,* opens a discussion on how things (e.g. technologies, artefacts) play an active role in shaping human morality and experience. More specifically, Verbeek (2011) introduces the notion of moral mediation to explore how technologies shape moral conduct by mediating concrete experiences and practices. By setting the prenatal ultrasound scan example, Verbeek (2011) explains that this technology presents the fetus in a novel way, altering how it is perceived, which in turn introduces several ethical issues, new forms of responsibility, and a reconfiguration of moral practices and choices. Building on the ultrasound example, parents' first visual experience of their unborn child can happen several months before

labour, allowing to examine medical abnormalities and inform future parents. The unborn child is framed as a potential patient, during the sonogram procedure, while parents have already been placed in a position to make a moral decision on whether they would proceed with the ultrasound scan since its availability. In these terms, a certain sense of responsibility is being shaped because of the potential of such a technology to identify risky labour. Building on this premise, Verbeek argues for a shift in understanding technology from being mere instruments to active moral entities.

These examples of technomorality and moral mediation are particularly helpful for understanding the role of AI in educational contexts. As with social media and prenatal ultrasound technologies, AI artefacts in higher education do not merely support or enhance existing practices but actively shape the conditions under which moral decisions are made. Technologies such as AGS and GAI chatbots introduce new possibilities for action while simultaneously foreclosing others, thereby reshaping how responsibilities, expectations, and ethical considerations are configured within educational practice. The concepts of technomorality and moral mediation thus provide an analytical lens for examining how ethical dimensions of teaching and assessment are co-constituted through human–technology relations. Rather than locating moral responsibility solely in human actors, this perspective foregrounds how AI technologies participate in shaping moral practices in higher education.

Concluding this chapter, the theoretical perspectives discussed above – postphenomenology, technological mediation, technopractices and technoperformances, and technomorality – provide a coherent framework for analysing the role of AI in higher education as a situated, mediated, and morally consequential phenomenon. Rather than approaching AI as a neutral tool or as an autonomous agent, this thesis adopts a relational perspective that attends to how educational practices, experiences, and moral responsibilities emerge through ongoing human-technology-world relations. This theoretical orientation informs my analytical focus on teachers' engagements with AGS and GAI chatbots, guiding attention to how these technologies mediate practices, reshape roles, and open up 'new' ethical questions within contemporary higher education.

# Methodology

This chapter outlines the research strategy and methods I adopted to approach the aim of the thesis. The chapter begins by offering an overview of the thesis' overall research strategy and rationale of methodological choices. The following sections describe the methods used in the studies, including a detailed overview of the data collection and analysis methods. Finally, the last section presents the ethical considerations of the research, including my reflections on research positionality.

This compilation thesis consists of four studies (Study I, II, III, and IV) that together examine how emerging AI technologies mediate higher education practices. Focusing on automated grading systems (AGS) and generative AI (GAI) chatbot technologies, I examine university teachers' perspectives and the ethical issues arising from such mediations. Guided by a postphenomenological perspective and technology mediation theory, the research strategy is plural and layered, in methods and analytical levels. Rather than treating AI technologies as neutral tools or as autonomous determinants of educational practice, the studies approach AGS and GAI chatbots as mediating technologies whose meanings, functions, and ethical implications emerge relationally in teachers' engagements with them. This theoretical orientation required a layered analysis exploring AI mediations from an experiential, perceptual and ethical dimension. These dimensions or levels of analysis facilited a cumulative building of understanding of how AGS and GAI chatbots' mediations were encountered and interpreted (*experiential* level), influenced and led to changes of how teachers perceive higher education practices (*perceptual* level) and had an ethical impact in their practices (*ethical* level). Additionally, this orientation necessitates a research strategy that combines conceptual and empirical approaches and attends to higher education practices and to teachers' expectations, experiences, and perceptions of technological emergence.

To address the research questions, I employed four complementary studies, combining conceptual and empirical methodologies. This approach allowed me to first engage deeply with the literature on AGS and examine the ethical issues related to their design and development. Building on this conceptual foundation, I then empirically explored how AGS are designed and developed – particularly in terms of the expectations surrounding them – as well as, how they mediate teachers' assessment practices. Similarly, the empirical studies on GAI chatbots provided a layered understanding of teachers' experiences

and perceptions of technological mediation in their practices, drawing on both experimental and qualitative approaches.

More specifically, to examine AGS and how they mediate higher education practices in relation to the assessment of student work, I conducted:

- Study I (Farazouli, 2024), which is a conceptual study examining the ethical considerations of AGS, reviewing the literature on AGS and analysing their specificities through a relational ethics approach; and
- Study II (Figueras, Farazouli et al., 2025), which is an interview study examining how AGS mediate university teachers' assessment practices. Interviews were conducted with AGS developers who are also university teachers using these systems, focusing on their expectations, experiences and disruptions AGS introduce into their practices.

Additionally, to examine GAI chatbots and how they mediate higher education practices in relation to teaching, learning and assessment practices, I conducted:

- Study III (Farazouli et al., 2024), which is an experimental study examining how GAI chatbots mediate teachers' perceptions of their assessment practices, though a Turing-test inspired experiment followed by focus group interviews with university teachers; and
- Study IV (Farazouli et al., 2025), which is a focus group interview study with university teachers examining how they experience and interpret the emergence of GAI chatbots and how GAI chatbots mediate their perceptions of their roles and practices.

Studies III and IV shared the data collection procedure, divided into a two-stage data collection plan, including an experiment phase followed by focus group interviews. Study III's core focus lies on the experiment and the post-experiment reflections of the participants, while Study IV focuses solely on the focus group interviews that occurred after we concluded the experiment and reflections part. Therefore, even though both studies shared the same data collection procedures, they included different parts of the interview data, differentiated in the analytical methods and were led by different research questions.

Such a research strategy reflects an understanding of AI technologies in higher education as emerging and relational phenomena whose roles, meanings, and ethical implications are not fixed in advance but are continually negotiated within practice. Rather than seeking to stabilise AGS and GAI chatbots as clearly bounded objects of analysis, the studies attend to how these technologies come to matter through teachers' engagements with them, including moments of uncertainty and disruption. The four studies offer conceptually informed and empirically grounded insights into how AGS and GAI

chatbots are experienced, anticipated, and negotiated by university teachers as part of their practices. By combining conceptual analysis with complementary empirical designs, the thesis maintains a methodological openness that mirrors the indeterminacy of the technologies under study, while remaining attentive to how educational practices are reconfigured through ongoing human–technology relations.

## Conceptual analysis

Study I explored the ethical considerations of AI automation in assessment, instantiating AGS as an emerging AI technology. This conceptual analysis of AGS and AGS literature draws on relational ethics and examines AGS' specificities regarding data, algorithms, and assessment, and the context in which these systems are used, including situations and purposes, actors and relations, time and place. More specifically, a relational approach to ethics is adopted to examine the fundamental ethical concerns and implicit assumptions that drive the development and deployment of automation in assessment. This work deliberately refrains from offering prescriptive recommendations or definitive solutions to the ethical challenges that such systems may raise in educational contexts. Rather, it argues for rethinking the ethics of AIED, and more specifically, the ethical considerations imbued in the design and use of AGS, from a relational standpoint, focusing on context, including situations and purposes, actors and their relationships, time, and place.

Study I adopts a conceptual methodology. The aim is to advance understanding of the ethical issues that arise with AGS, rather than examining empirical data. Conceptual research is widely recognised as a legitimate form of academic work, particularly in fields that are still in a process of consolidation. Jaakkola (2020) argues that conceptual contributions are necessary for the development of new perspectives, models, or frameworks that can then inform empirical inquiry. Aligning with Jaakola's types of research designs of conceptual work, Study I could be characterised as a *theory adaptation* article, that aims to revise the current understanding of what AGS *is* and *do* when deployed in educational settings, problematise the existing approaches to AI ethics literature and introduce an alternative aproach in examining ethical issues, from a relational ethics approach, of AI artefacts, such as AGS. As Jaakola (2020) identified, similar examples of such type are the work of Alexander et al., (2023) and Hillebrand et al. (2015).

The methodological process followed in this study can be described in three interrelated steps. In the first stage of conducting the study, I got familiar with AGS and its variations in terms of architecture and design, algorithms and technical features, as well as functions and purposes, by drawing on the first-hand experience, being part of a research group developing such systems, and conducting a narrative literature review in a body of research published in the

last two decades. AI-powered AGS is a relatively new research topic, and the area is still not well-developed; therefore, a narrative review was considered appropriate because the purpose of this research was to provide a descriptive overview and understanding of what AGS are, how they function, and the contexts in which they are applied. A narrative review offers a way to summarise existing knowledge on a topic while also providing a critical and/or interpretive perspective on the body of literature as a whole. Narrative reviews enable researchers to highlight the current state of the field, point to areas for further development, and introduce alternative viewpoints or theoretical considerations (Sukhera, 2022). For this reason, narrative reviews are valuable not only in emerging areas where research is limited but also in well-established fields, where they can generate fresh insights and new ways of framing existing evidence (Rumrill & Fitzgerald, 2001; Sukhera, 2022). Unlike systematic reviews that follow standardised protocols for search and selection of literature, conducting a narrative review on AGS allowed me to explore a wider breadth of sources (Green et al., 2006), including research articles, short papers, workshop papers, posters, conference proceedings, as the inclusion criteria were deliberately broad: any source that provided definitions, descriptions, analyses, or evaluations of AGS was considered relevant. I made several searches over time in IEEE Xplore, Scopus, Google Scholar, and ACM Digital Library, trying to follow publications, conferences and workshops in the field of machine learning, AI, natural language processing and AIED, using the following string:

> ((“Automated assessment” OR “automatic assessment” OR “automated evaluation” OR “automatic evaluation” OR “Automated grading” OR “automatic grading” OR “Automated scoring” OR “automatic scoring” OR “automated grading” OR “automatic grading” OR “ automated marking” OR “automatic marking” OR “ automated testing” OR “automatic testing” OR “Automated essay scoring” OR “Automated evaluation system” OR “Automatic writing evaluation” ) AND ("Higher education" OR "tertiary education" OR “college” OR “Third level education”))

In the second stage of conducting the study, I also reviewed literature on ethics of AI, learning analytics and AIED published since 2000 as well as policy documents, such as guidelines and codes of conduct, issued by national and international bodies (i.e.: *Ethics Guidelines for Trustworthy AI* by EU, *The OECD Artificial Intelligence Principles*, *Recommendation on the ethics of artificial intelligence* by UNESCO, the Beijing Academy of Artificial Intelligence (BAAI) *AI Principles*) and secondary reviews of published research on ethics and AI. At this stage, I aimed to identify how ethics is addressed in relation to AGS or other AI technologies for assessment. As a result of such a

review, I identified the gap in ethics literature not addressing AGS as a technology and inductively identified the major ethical issues that relate to automated educational decision making and AGS.

In the third stage, drawing on relational ethics literature and inspired by the structure Holmes et al.'s (2022) suggest to unpack AIED ethical issues, I introduced and applied relational ethics as the central conceptual lens to analyse and present the ethical considerations of AGS. Relational ethics, as articulated by Bergum and Dossetor (2005) and developed further in education and technology research (Birhane, 2021; Decuypere & Simons, 2016), prioritises relationships, interdependencies, and situated contexts. Through this perspective, I reinterpreted ethical issues of AGS across three domains-data, algorithms, and assessment practices – highlighting the power dynamics, agency, and vulnerabilities that rational approaches tend to obscure.

## Interviews

Study II is an interview study exploring the practical implementation of AGS by teachers who also serve as developers of these systems. By engaging in interviews, the participants shared details about their tools, unfolding how they initiated and developed AGS, and what kind of challenges they encountered when using them in practice. In-depth semi-structured interviews were conducted with participants who held multiple roles as teachers, researchers, and developers in the development and deployment of AGS. The interviews took place within the context of a Swedish public university. The interviews were conducted one-to-one, in either the participants' workplace or via Zoom, based on the interviewees' preference, from mid-November 2022 to mid-January 2023, each lasting one hour on average.

A purposive sampling approach was employed to select participants who met two criteria: a) involved in the design and development of an AGS, and b) deployed AGS into their practices or followed the deployment of AGS in assessment practices. All participants in this study were employed within the same Computer Science department of a Swedish university at the time of the interviews. The participant group consisted of six males and one female, with varying ages, levels of experience, and academic positions (Table 5). While the relatively small sample size may limit the generalisability of the findings, the depth and expertise of the interviewees offer valuable and nuanced perspectives within the scope of the study. I was also mindful of the gender imbalance in the participant pool, as all but one of the interviewees were men, reflecting the composition of the teachers/developer population in the department. This gender disparity represents a limitation in terms of representation. However, it is important to note that this study did not specifically compare or analyse differences between genders.

*Table 5 Participants Study II*

| Code ID | Role in AGS (researcher/developer/teacher) | Position | Years in Academia |
|---|---|---|---|
| *I-1* | **D-T** | University Lecturer | 25 |
| *I-2* | **R-D-T** | Assistant Professor | 10 |
| *I-3* | **R-D-T** | Research & Teaching Assistant | 8 |
| *I-4* | **D-T** | Associate Professor | 35 |
| *I-5* | **R-D** | Research Assistant | 1 |
| *I-6* | **R-D** | Research Assistant | 1 |
| *I-7* | **R-D-T** | Professor | 15 |

To facilitate the interviews, an open-ended and semi-structured interview guide was designed. This guide allowed participants to share their experiences, express their concerns, and discuss their perspectives on AGS. The interview guide covered various aspects, including the design process, the rationale and purpose behind the development of the specific AGS they were involved with, and inquiries about any concerns, risks, or ethical challenges they encountered while using such systems. The interviews were structured into three parts to ensure a comprehensive exploration of the topics at hand:

- First, introductory questions were asked to gather information about the backgrounds and current positions of the interviewees.
- The focus then shifted to AGS, and participants were asked questions related to the AGS they developed or used. These questions included requesting a description of the system, its aim and purpose, the interviewees' role in the development process, intended use, and functionality.
- The final part of the interview delved into participants' work practices with the AGS, ethical considerations, and the impact of the tool on student-teacher relationships. Specifically, several topics were discussed, such as obtaining student consent for data collection and de-

ploying AGS for assessment purposes, technical and ethical challenges encountered during development and implementation, fairness considerations, and the potential misuse of the tool.

The data were analysed using the reflexive thematic analysis approach outlined by Braun and Clarke (2019, 2021). The analysis was iterative, involving multiple cycles of theme identification and refinement. The first two authors co-led the analysis, alternating between in-depth analysis sessions and presenting/discussing the results with the other authors. Adopting an exploratory approach, we first familiarised ourselves with the empirical material by separately listening to the audio recording, conducting close readings of the transcripts, and noting down meaningful units and interesting sections (i.e., parts that sparked our surprise, curiosity or appeared worthy of further exploration). Subsequently, open coding, forming initial themes, was performed separately by the first two authors, who then met several times to share our reflections and consolidate common themes. These themes were further discussed with the rest of the research team, crafting the final themes as presented in Tables 6 and 7.

*Table 6 Study II themes - Improve assessment & Manage workload*

| **Theme** | **Improve assessment process** | | | |
|---|---|---|---|---|
| **Categories** | Course evaluation | Track students' learning | Fair/unbiased grading | Advance students' learning |
| **Codes** | analysis of the patterns of failure<br><br>difficult concepts | easier to track the progress<br><br>identify patterns<br><br>opportunity for continuous assessment | stable and more consistent among teachers<br><br>equal treatment among students<br><br>"more objective"<br><br>reduce human bias-fatigue hunger | real programming environment<br><br>training function-more opportunities to experiment and test answers<br><br>earlier feedback -students' autonomy and confidence |

| **Theme** | **Manage Workload** | | |
|---|---|---|---|
| **Categories** | Support in managing and organizing assessments in smaller parts | Speed up the assessment process | Allocation of time to other teaching activities |
| **Codes** | Managing large courses and tight deadlines<br><br>Consistent grading | Saving time<br><br>AGS as compiler -testing functionality of the-code/program- and checking for technical and syntactical issues in the code | More time on other teaching responsibilities preparing lectures revising syllabi<br><br>conducting course evaluations<br><br>developing plans for subsequent course improvements |

*Table 7 Study II themes – AGS breakages*

| **Theme** | **AGS breakages** | | | | | |
|---|---|---|---|---|---|---|
| **Categories** | Assessment design | Bias and fairness | Limited space for creativity and complex tasks | Students' discouragement | Student-teacher relations | Transferability issues |
| **Codes** | Difficulty to create Qs that fit AGS<br><br>Hard to modify and update assessment design | biases in training processes<br><br>limited diversity in data annotators<br><br>bias on historical student data<br><br>unfair and discriminatory assessment by the system | AGS cannot assess code quality<br><br>AGS limitation to assess free text responses<br><br>AGS limitation to assess problem solving tasks<br><br>limitation in students' motivation<br><br>oversimplification of the assignment | Students' potential stress<br><br>Students' dropout<br><br>students less keen on submitting carefully<br><br>students less interested in receiving feedback by the systems<br><br>instant and low-quality feedback (compiler 'error')<br><br>AGS not actively encouraging student reflection | disruptions in their communication and interaction through assessment<br><br>student-teacher communication loss<br><br>loss of personal connection<br><br>lack of nuanced insights gained from manual exam reviews | difficulty to replicate<br><br>AGS tailored to each course<br><br>AGS tailored by each examiner<br><br>AGS trained on specific student data |

## Experiment

Study III, explored the mediating role of GAI chatbots in university teachers' assessment practices, focusing on the humanities, social sciences and law studies. Using ChatGPT as an indicative example of an advanced GAI chatbot, a Turing test-inspired experiment on assessment was designed and conducted, where university teachers were asked to assess student and bot-written responses to exam questions.

The Turing Test (TT) was established in 1950 by Alan Turing as a method of testing whether a 'machine can think'. However, several variations of the test have been practised by researchers, such as the Reversed Turing Test (Mayumi Takaya et al., 2013). The design of our experiment is mostly inspired by TT and the Feigenbaum Test (Feigenbaum, 2003), which is an alternative test of TT where a human expert in a field is called to assess whether a response is generated by a human or a computer. This design of experiment was selected in this study as I aimed to examine teachers' experiences of assessment in the potential presence of ChatGPT. More in particular, such a design allowed me to study the GAI chatbot-mediated role when teachers assessed the texts (responses to examinations) we asked them to.

With support by the research team, I conducted four experiments where the participants from each department were asked to assess five to six responses to a home exam from an introductory course at the undergraduate level in their department. Three of the texts were student responses previously graded as excellent, good and adequate. More specifically, we asked the course leaders from each department to share with us indicative cases of student responses previously graded as A, C and E, in order to ensure a variety in the quality of the texts meeting the grading criteria. Two of the participants who collaborated with us in recruiting teachers and validating the material used in the experiment participated only in the focus group interviews. The participants (n=22) were given a grading sheet, a description of the assignment and six responses.

Moreover, three ChatGPT-generated texts were used, which I had previously manipulated at different levels. More specifically, I conducted three levels of prompt engineering, from using verbatim the questions described in the exams to crafting prompts for the chatbot, guiding it to meet the grading criteria. The first category of ChatGPT texts used in our experiments referred to verbatim ChatGPT output, where the description of the exam question was used as a prompt. Hereafter, I will refer to this type of text as ChatGPT0. To generate the second category of ChatGPT texts, I used as a prompt to the chatbot the English translation of the description of the exam question when it was written in Swedish (philosophy, law, sociology, education) and asked to add references to the literature. Hereafter, I will refer to this type of text as ChatGPT1. Finally, the third category of ChatGPT texts refers to a combination of multiple ChatGPT outputs. To generate these outputs, I asked for

longer texts referring to the minimum and maximum word limit of the exam along with specific citations and references linked to the course literature. In addition, I crafted different prompts where parts of the assignment were run separately. Finally, this third category of ChatGPT-generated texts hereafter will be referred to as ChatGPT2. All three types of texts (ChatGPT0-2) were transferred to Word files and followed the format instructions described in the exams.

I wanted to include three variants of chatbot texts in the experiment to examine how convincingly ChatGPT can generate responses at different levels of prompt engineering and output manipulation, with varying degrees of human involvement. The director of studies from each department had approved all ChatGPT texts used in the experiments as relevant and adequate responses to assignments, except for the ChatGPT0 response for the home exam in the Department of Education, where it was suggested that the text was of very low quality and, therefore, was not included it in the TT experiment.

Finally, although the participants have been informed about the research aim and focus on AI and could assume that some text could be chatbot-written, they were asked to assess the quality of the texts and assign a score and not identify the chatbot- and student-written ones. Hence, the teachers did not know the potential number or which were the chatbot-generated responses and therefore, the texts were given to all the participants in the same but random order, blending the student and ChatGPT responses. More specifically, the order we used is: 1. student text previously graded with A (Text A), 2. ChatGPT1, 3. student text previously graded with C (Text C), 4. ChatGPT0, 5. ChatGPT2 and 6. student text previously graded with E (Text E)[14].

With regard to the sampling of the study, 24 university teachers (P1-P24) were included from four different departments in a Swedish University (Table 8). The participants were recruited in collaboration with the directors of studies from each department and formed four groups of four to eight teachers. 10 participants had previously used ChatGPT or similar chatbots, while the rest 14 had only read about such tools and seen several examples shared in the media or by other colleagues. The main sources of data were grading sheets completed by the participants, where they assigned grades to the texts following each department's grading format and the verbatim transcripts of the four focus group interviews. Additionally, field notes during the experiment and the interviews, along with participants' notes on the texts they assessed, were used as supporting material for validating and triangulating the data.

[14] All ChatGPT-generated responses were checked for plagiarism using Stockholm Univeristy's detection software, indicating that only 4 out of 14 ChatGPT texts were flagged as plagiarised (5-11%), located in the references lists, repetitions of the question and references to factual knowledge or established theory, which would not be considered plagiarised text in a realistic assessment setting.

*Table 8 Participants Study III & IV*

| Demographic variable | Level of variable | n = 24 |
|---|---|---|
| ***Gender***[1] | Female (%) | 10 (41.7) |
| | Male (%) | 14 (58.3) |
| ***Age (yr)*** | Average | 42.7 |
| | Min-Max range | 23–69 |
| ***Faculty*** | Humanities (%) | 8 (33.3) |
| | Law (%) | 6 (25) |
| | Social Sciences (%) | 10 (41.7) |
| ***Department*** | Philosophy (%) | 8 (33.3) |
| | Law (%) | 6 (25) |
| | Education (%) | 4 (17.7) |
| | Sociology (%) | 6 (25) |
| ***Academic position*** | Ph.D. Student (%) | 6 (25) |
| | Assistant Professor (%) | 12 (50) |
| | Professor (%) | 6 (25) |
| ***Years as university teacher*** | Average | 13.4 |
| | Min-Max range | 1-36 |

[1] For the description of the gender, participants were asked to freely describe how they identify their gender.

## Focus groups

Following the TT experiments, semi-structured focus group interviews were conducted with the participants from each department, divided into two distinct sessions, each serving a different purpose. The first moment, which lasted 45 to 60 minutes, included the data collection of Study III, aiming to reach a deeper understanding of how the teachers assessed the text and how they reasoned about their assessments. More specifically, during that moment of data collection, the participants collectively shared and discussed their reflections on the texts they read. The second moment began after closing the post-experiment reflections phase, lasted 90 minutes and included the data collection of Study IV. During that moment of the focus group interviews, the participants were asked to discuss how they experience and perceive the emergence of GAI technologies in relation to their practices as well as how they reflect on

the potential impact of GAI chatbots in their context of teaching and assessment practices. I, accompanied by another researcher from the research team, were present to moderate the focus group interviews. The interviews were audio recorded and transcribed, and then analysed by me in consultation with the other authors.

The strategy for analysing the qualitative data of both studies was inductive and greatly inspired by Graneheim and Lundman's (2004) work, where the texts (in this case transcriptions) include multiple meanings and interpretations of them by the researchers are unavoidable. Therefore, the same authors suggest that in order for qualitative research to achieve a high level of trustworthiness in findings, there has to be a systematic way of creating a hierarchy of concepts, such as meaning units, codes, categories and themes (Graneheim & Lundman, 2004). Following this paradigm of qualitative content analysis, after I had obtained a sense of the whole by reading the transcripts, I divided the texts into meaning units where participants referred to their assessments during the TT-inspired experiment. These meaning units formed condensed meaning units and, through abstraction, created several codes which were then grouped into thematic categories and finally formed our main themes. The themes from both studies are presented in the tables below.

In Study III, which examined ChatGPT's performance in home examinations and assignments, the identified themes concerned the quality of the generated texts, including their perceived strengths and weaknesses (Table 9), and the textual features that participants perceived as indicators of AI-generated rather than student-written texts (Table 10).

*Table 9 Study III themes – Chatbot performance in home exams / assignments – quality of texts*

| **Chatbot performance in home-exams/assignments (a)** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Theme** | Quality of texts | | | | | | |
| **Category** | *Strengths* | | | | *Weaknesses* | | |
| **Sub-category** | Good language | Precision | Consistency | Creativity | Argumentation | References | Relevance |
| **Codes** | Good structure<br>No typos<br>Good grammar | Answers the qs "all points are there" | Short but specific<br>Lists of arguments | Innovative responses<br>Unusual responses | Weak or no arguments | No engagement with course literature | Irrelevant content<br>Lack of examples<br>Difficulty to engage with given context |

*Table 10 Study III themes – Chatbot performance in home exams/assignments – suspecting AI-generated text*

| **Chatbot performance in home-exams/assignments (b)** | | | | |
|---|---|---|---|---|
| **Theme** | Suspecting AI-generated Text | | | |
| **Category** | *Weird text* | | *Impersonal* | *Resemblance* |
| **Sub-category** | Non-human | Human | Inconclusive | |
| **Codes** | Non-sense statement<br>Weird words (translations, synonyms)<br>CGPT writing style<br>Repetition of the prompt | Not terminology used in class<br>Repetition in the arguments<br>No references | Vague arguments<br>No point of view<br>Empty statements<br>Too simple | Similarity between the texts |

Additionally, Study III examined ChatGPT's mediating role in teachers' assessment practices. Table 11 outlines the categories, subcategories, and codes that illustrate the pattern of heightened criticality, including tendencies toward downgrading, suspicion of AI-generated texts, and raising grading standards.

*Table 11 Study III themes – Chatbot mediating role on teachers' assessments - teachers were more critical when assessing student texts*

| **Chatbot mediating role on teachers' assessments** | | | | | | |
|---|---|---|---|---|---|---|
| **Theme** | Teachers Were More Critical When Assessing Student Texts | | | | | |
| **Category** | *Downgrading in most cases* | | | *Suspecting bot-written text (falsely)* | | |
| **Sub-category** | Rarely put a high grade | **Not accepting simple flaws** | Higher standards for a pass | "too good effect" | Not connected to the context | Chatbot mistakes |
| **Codes** | % high grades | Structure<br>Coherency<br>References<br>Length | % failed | Too good language (flawless, no typos, no grammar mistakes)<br>Too good content (rich, accurate) | Engagement with irrelevant literature not listed in the course literature | Inaccurate content<br>Irrelevant content |

*Table 12 Study III sub-category – Not accepting simple flaws*

**Not accepting simple flaws**

| Meaning Unit | Condensed Meaning Unit | Code (Sub-category) |
|---|---|---|
| Repetition | Structure | Simple flaws in student responses not directly connected to the home-exam/assignment learning objective |
| Mixing accurate and inaccurate content | Coherency | |
| Lack or wrong citations | References | |
| Inaccurate (made-up references) | | |
| Engagement with relevant literature not listed in the course literature | | |
| Short answers | Length | |

In Study IV, examining teachers' experience of the emergence of GAI, three main themes were identified regarding teachers experiencing the emergence of GAI as an alarming and overwhelming situation, as being vulnerable and as invoking several ethical issues (Table 13).

*Table 13 Study IV themes – Experiences of the emergence of GAI*

| *Experiences of the emergence of GAI* | |
|---|---|
| **Theme:** Alarming & overwhelming situation | |
| **Codes** | **Condensed meaning units** |
| *Alarmed by the hype* | Stream of news on cheating;<br>Reactions from big universities |
| *Necessity for fundamental changes* | Misfit of the current educational system to GAI;<br>Rethink defining knowledge and performance |

| **Theme:** A state of vulnerability | | |
|---|---|---|
| **Codes** | **Condensed meaning units** | |
| *Vulnerable to mistakes* | Uncertain about how to act;<br>Limited knowledge of GAI;<br>Limited exposure to using GAI | |
| *Low confidence* | Sense of incompetence;<br>Discouraging for teachers with little experience | |
| *Overwhelming change* | Too fast and sudden technology;<br>Worry about not being ahead of students | |
| **Theme:** Emergence of ethical challenges | | |
| **Categories** | **Codes** | **Condensed meaning units** |
| Unequal access to the tools | *Limited access/equity issues* | Paid versions of tools;<br>Unequal access to GAI tools |
| GAI biases | *Bias* | Biased content of GAI tools |
| | *Inaccuracies* | Inaccurate content of GAI tools |
| Treating students fairly /Fairness | *Balancing criticality* | Falsely accusing students |
| | *Fair treatment* | Weakening trust in students;<br>Students cheating when using GAI tools |

Furthermore, Study IV's themes regarding teachers' perceptions of the potential transformation in their practices centred around rethinking assessment, reevaluating teaching priorities and perceiving student learning as being at risk (Table 14).

*Table 14 Study IV themes – Perceptions of the potential transformation in teaching and learning practices*

| *Perceptions of the potential transformation in teaching and learning practices* | | |
|---|---|---|
| **Theme:** Rethinking assessment | | |
| **Codes** | **Condensed meaning units** | |
| *Shifting to traditional assessment methods* | Banning home exams;<br>Sit-in exams;<br>Oral exams | |
| *Exploring alternative types of assessment* | Multiple forms of assessment;<br>Incremental submissions | |
| *Increasing transparency in AI use in assessment* | Honour statement of non-use;<br>Declaration of use | |
| **Theme:** Re-evaluating teaching priorities | | |
| **Codes** | **Condensed meaning units** | |
| *Critical thinking* | Critical on resources;<br>Evaluative judgement | |
| *Think independently* | Confidence and ownership of knowledge | |
| *Ethical use of GAI* | Authorship and ownership;<br>Academic integrity | |
| *Discovering the pedagogical value of GAI* | Summarising and explaining long/difficult texts;<br>Working on advanced/creative tasks;<br>Focus on critical reading | |
| **Theme:** Student learning at risk | | |
| **Categories** | **Codes** | **Condensed meaning units** |
| Weakening students' skills | *Automation and deterioration of generic competencies* | Writing skills;<br>Reasoning skills;<br>(re)searching skills;<br>Responsiveness in real-world situations |
| Desocialisation of learning | *Learning process* | Harming social construction of knowledge;<br>Working less with peers |
| | *Learning content* | Missing the sociality within learning content |

## Research ethics

Conducting research in the social sciences that involves empirical data inevitably entails a certain form of power – particularly in representing participants' views and the social situations in which I, as a researcher, have engaged them. Recognising this, I have been attentive to how the process of data collection, interpretation, and representation could shape the voices and experiences of those involved in this study. Across the empirical studies included in this thesis (Studies II–IV), I have reflected on the potential social and emotional impact that the research data collection moments (individual interviews, experiments, and focus groups) might have had on participants. These moments functioned not only as spaces for data collection but also as opportunities for participants to engage in reflection. During each data collection session, participants were invited to "pause and share" their reflections on their AI-mediated teaching and assessment practices. This process often prompted them to articulate their reasoning for adopting or avoiding AI (especially in the case of AGS) and to consider the broader implications of these technologies in their professional contexts. Such reflective moments eventually led, on certain occasions, to expressions of vulnerability, feelings of uncertainty, being mistaken, or being limited in their teaching practices. Throughout the research process, care was taken to create an environment in which participants felt respected, safe, and comfortable to share their experiences. To ensure this, the research team invited the participants to choose the time and format of participation (in person or online) according to their preferences and convenience. Furthermore, before each session, we reviewed the consent form and answered any of their questions to ensure that participants fully understood confidentiality procedures and their rights, including the option to decline questions or withdraw at any point without consequence. In order to foster mutual trust, minimise potential discomfort, and support an open and respectful exchange of perspectives, the interviews and group discussions were conducted in a conversational and non-evaluative manner, allowing time for pauses, clarification, and reflection.

This thesis follows the Swedish Research Council (2024) *Good Research Practice* guidelines as well as the European code of conduct for research integrity (ALLEA, 2023), based on the principles of reliability, honesty, respect and accountability. More specifically, with regards to *reliability* and *honesty* of my research findings, I have ensured a high level of transparency for following the research process in terms of research gap/problem formulation, theoretical and epistemological positioning, methodological choices as well as presentation of the analysis and research findings, by providing details of the steps I followed, the rationale as well as the analysis schemes I base the

findings of the thesis. In relation to the principle of *respect,* data collection processes were conducted with respect to the participants' time and space preferences, ensuring a comfortable and trusted environment for them to be part of. Participants were informed about their right to withdraw from the research at any stage without any consequence or explanation required. Additionally, regarding *accountability*, I have continuously sought support and feedback from academic and professional communities, including my institutional research environment (Department of Education, Stockholm University), the Higher Education Learning Practices (HELP) research group, the Swedish Higher Education Research Network (SHERN), and the WASP-HS project and graduate school. Engagement in these networks has provided opportunities for ongoing reflection on research ethics, integrity, and the broader societal implications of my work.

Ethical integrity was a central consideration in the design and conduct of this thesis. Prior to commencing data collection, ethical approval was obtained from the Swedish Ethical Review Board (Dnr 2022-0105-01), and all procedures adhered to Stockholm University's research integrity and ethics policy. The empirical material was audio-recorded, transcribed verbatim, and anonymised to ensure confidentiality. All the data was handled and stored in accordance with the Declaration of Helsinki, the General Data Protection Regulation (GDPR) and Swedish Authorities. All participants were informed about my research project and its goals and asked to sign consent forms about participating. Additionally, they were informed that their participation is completely voluntary and they can opt out at any time without specifying the reasons. Participants were recruited through email invitations, which included information about the research's main objective, its affiliation with a research project 'Ethical and Legal Challenges in Relationship to AI-Driven Practices in Higher Education'[15], and approval from the corresponding Ethics Review Authority. Once participants agreed to participate and scheduled a meeting time, they received a confirmation email containing an informed consent form, which they were asked to read and sign prior to the interview or experiment.

Finally, while this research aligns with Stockholm University's Open Data Strategy, full open sharing of data was not possible due to the sensitive nature of the material, which included participants' personal and philosophical reflections on their professional practices. Instead, I ensured transparency through detailed data management documentation, developed using Stockholm University's Data Management Plan tool. Supplemental materials, including the interview guide, coding of data, consent forms, along with all documents shared with participants of the studies and materials used in the studies, are available upon request.

---

15 https://wasp-hs.org/project/ethical-and-legal-challenges-in-relationship-to-ai-driven-practices-in-higher-education/ (Retrieved February 26, 2026)

## Notes on the researcher's positionality

Who I am as a researcher is inseparable from my background. I come to this work as a critical AI female researcher and educator, with professional and academic roots in computer science and education, experience in teaching across contexts from early childhood to professional training, and a philosophical curiosity about how ethics takes shape in practice. These strands of my identity shape how I approach this project and what I question: I pay attention to the 'technologies themselves', I foreground the voices of educators as central stakeholders in education, and I unpack the ethical issues that arise in educational practice. During my studies, I developed a critical orientation toward the intersections of technology and education, becoming increasingly interested not only in *what* and *how* things happen but in *why* they happen, according to *whom*, and with what *implications*.

Inspired by Pillow's (2003) notion of *uncomfortable reflexivity*, who sees reflexivity not as a methodological tool for acknowledging position, power and view, aiming to sanitise and imply neatness or moral certainty, but rather as a moment to recognise and value research that shows the struggle itself, including imperfection, uncertainty, and ongoing negotiation that characterise researcher inquiry, I view reflexive practice not as a methodological reassurance or a means to perfect representation, but as an ongoing process of accountability to participants' views for self-determination and practices, and to my own entanglement in the ethical and epistemic tensions of this research. In practice, this stance was enacted by maintaining openness to uncertainty throughout the research process. I, together with the research team involved in the studies included in this thesis, documented interpretive tensions during data analysis, discussed divergent readings of participants' accounts, and intentionally retained excerpts that did not neatly fit the emerging themes. Such practices aimed to resist the temptation of closure and instead foreground the complexity, contradiction, and ambiguity that accompany qualitative inquiry. Additionally, rather than framing the findings of this thesis as a finished end, I recognise it as part of my ongoing development as a researcher and inquiry in the area of how technologies mediate higher education practices. Through this approach, reflexivity is not treated as a discrete methodological stage, but as a continuous, accountable engagement with the ethical and epistemic entanglements of studying AI mediations in teachers' practices and the ethical issues emerging from such mediations.

In relation to both the topic of study and the participants, I occupy a position that is partly *insider* and partly *outsider*. I am an insider in the broader sense of being a colleague within Swedish higher education, as a member of the teaching staff and as a former AI system developer, sharing certain professional languages and values. Simultaneously, I am an outsider

in that I have not been directly involved in the participants' specific pedagogical practices or in their interpretations of technological artefacts. This dual position, which Dwyer and Buckle (2009) describe as being "the space between", requires continuous reflexivity to recognise how proximity and distance each shape interpretation, representation, and ethical responsibility. This dual position was also continuously negotiated throughout the research process. I drew on my insider knowledge of higher education to establish trust and frame discussions in ways that resonated with participants' professional realities. At the same time, reflexive discussions within the team, along with analytic memos, were used to document moments when familiarity risked narrowing interpretation and to foreground alternative readings. In this way, the insider/outsider position became a methodological resource, allowing for both empathetic engagement and critical interrogation of the assumptions embedded in participants' and researchers' perspectives alike.

# Summaries and findings of the Studies I-IV

In this chapter, I present the extended summaries of the Studies I, II, III and IV focusing on their main findings and conclusions.

## Study I: Automation and Assessment

Exploring Ethical Issues of Automated Grading Systems from a Relational Ethics Approach

Study I provides a critical examination of automated grading systems (AGS) by approaching ethics not as a matter of abstract principles but as something that emerges in the situated and relational practices of assessment. This perspective is valuable for understanding how AGS actively mediate relationships between students, teachers, and institutions, foregrounding ethical issues as context-dependent and evolving rather than universal or fixed.

This study adopts a critical standpoint, using a relational ethics approach, to discuss and unpack ethical issues of automation in assessment. Assessment as a 'socially situated interpretive act' (Shay, 2008, p. 309), lying at the core of academic activities, constitutes a situated practice embedded in teachers' identities and is dependent on systems' infrastructures and values. Looking at assessment practices as a vital aspect of educational systems, which is inevitably relational, this study focuses on automation to discuss ethical issues that can emerge when AGS are introduced to assessment practices.

In contrast to dominant ethics approaches reflected in broad ethical guidelines, a relational ethics approach helps to enact a situated and inclusive understanding of critical ethical and moral issues at stake when designing and using AGS in assessment practices. More specifically, dominant approaches to ethics tend to comprise close-ended guiding documents indicating overarching principles without recognising the ever-evolving character of technological progress as well as the ever-expanding web of relations in the context of education. In line with Birhane's (2021) argument that algorithmic systems, such as AGS, are not developed and deployed in a social, historical, and political vacuum, relational ethics highlights the importance of examining such automations in the context in which they are embedded. Accordingly, the context of the design and use of AGS in assessment practices should be viewed

as a particular and enacted result of the various interactions, negotiations, and relations that occur within it. Bringing forward values expressing relations and situating ethics in specific settings and contexts, this study argues for understanding ethical challenges in terms of the emerging relationships and attitudes that are amplified or reduced with AGS in educational settings.

In this study, I identify three dimensions regarding ethical issues surrounding AGS: data, algorithms, and assessment. In relation to data, questions arise around student autonomy and agency, particularly concerning consent, ownership, and the use of student work in training datasets – 'invisible labour' (Bearman et al. 2020; Cerratto Pargman et al. 2023). Furthermore, since AGS are developed within specific institutional and cultural contexts, issues of transferability and adaptability become apparent when such systems are applied elsewhere, often carrying forward contextual biases or losing validity over time. With regard to algorithms, the opacity of computational processes could make it difficult for educators and students to interpret or contest AGS decision outputs. Moreover, algorithms inevitably embody particular values and assumptions, raising concerns about the morality of computation and design and the ways in which certain assessment logics are privileged over others. The possibility of algorithmic injustice is also present, as misclassifications or systematic errors can unfairly affect students, with limited mechanisms for redress. Finally, the dimension of assessment highlights the risks of desocialisation, where assessment is reduced to technical pattern recognition and the dialogical and interpretive aspects of grading are weakened. Alongside this, there is the potential devaluation of teachers' expertise, as professional judgement is displaced or marginalised, even as new, less visible forms of oversight work emerge. Together, these issues illustrate that AGS not only introduce technical and procedural considerations but also reconfigure educational practices and relationships in ways that demand ongoing, situated, and relationally attuned ethical reflection.

The ethical issues explored in this study are not intended to represent an exhaustive account of all possible concerns associated with AGS and do not aim to function as a linear and static checklist of issues to be addressed. Rather, they form instances of issues that relational approaches to ethics might consider as an alternative path to exploring potential challenges and dilemmas, but also opportunities that AIED might introduce. With regard to future research orientations, this study suggests that both conceptual and empirical research aiming to examine the ethics of AGS and AIED could further explore ethical issues that arise in practice and shed light on AI systems' specificities in relation to a given setting.

## Study II: Promises and Breakages of Automated Grading Systems

A qualitative study in computer science education

Study II focuses on university staff with expertise in researching, developing and/or using AGS for assessment in higher education. Rather than examining AGS only as technical artefacts, the study positions them as socio-technical systems embedded in professional identities, institutional structures, and pedagogical goals. Firstly, this study explores the broader aspects or qualities surrounding AGS that the participants find valuable. This includes the specific characteristics of the AGS, their benefits for assessment practices, and their overall contribution to university teachers' professional practices. Secondly, inspired by the Science and Technology Studies (STS) concept of *breakages* (Jackson, 2014, 2016), the study examines the potential constraints and disruptions associated with the integration and mediations of AGS in assessment practices, which were then categorised as 'breakages'.

The interviews revealed what the participants found valuable about these systems and what they identified as problematic or limiting. On the side of perceived value, participants emphasised two major contributions of AGS. First, they were seen as important tools for workload management. By accelerating grading and automating repetitive tasks, AGS were portrayed by the participants as they could free teachers to focus on other pedagogical activities, such as mentoring and one-to-one student support. Second, participants highlighted the potential of AGS to improve the quality and consistency of assessment. Systems were praised for reducing human subjectivity, offering more uniform grading across large cohorts, and providing students with faster and sometimes more detailed feedback. The underlying notion of "mechanical objectivity" resonated strongly, with participants suggesting that AGS could produce fairer results by minimising the influence of individual biases.

At the same time, the study uncovered a range of constraints and unintended consequences, analysed through the STS concept of breakages (Jackson, 2014, 2016). Breakages are not failures in the narrow sense, but moments where systems encounter frictions, disruptions, or unforeseen effects that complicate their use. In the case of AGS, participants described several such breakages. Assessment design challenges were frequent: creating assignments that could be graded by AGS required substantial upfront effort, technical expertise, and continuous adjustment, often offsetting the time-saving benefits of automation. Transferability issues emerged when systems developed for one course or context were difficult to adapt elsewhere, limiting scalability. Concerns around fairness and bias also surfaced. While AGS were often presented as neutral, participants recognised that systems built on historical datasets risked reproducing the very inequities they sought to overcome. The

reliance on predefined categories and patterns also created blind spots, with AGS struggling to capture creativity, nuance, or unconventional approaches in student work. This contributed to another major breakage: the limited capacity of AGS to assess complex or qualitative aspects of assignments, particularly in open-ended tasks that demand interpretation or judgment. As a result, the participants stated that there was a danger of narrowing assessment practices, with assignments being redesigned to suit what the system could measure rather than what educators considered most pedagogically valuable. Another theme concerned the erosion of teacher–student communication. Participants noted that when AGS became central to assessment, opportunities for dialogical feedback and interpretive engagement between teachers and students were reduced. This "desocialisation" of assessment risked weakening the relational dimensions of learning, as feedback became more technical and less conversational. Participants also observed unintended effects on student practices, with some learners adapting their submissions to the perceived logic of AGS-optimising for what the machine could recognise rather than engaging deeply with the material. This shift raised concerns about surface learning, strategic compliance, and the loss of creativity in student work.

These findings underscore the ambivalent role of AGS as both promising and disruptive, potentially offering efficiency and consistency while simultaneously introducing new frictions and ethical dilemmas. By foregrounding both their promises and breakages, the study shows that AGS are not neutral tools but socio-technical systems that reshape assessment design, teaching roles, and student engagement. Recognising imperfections in AGS is crucial, as their integration can disrupt the relational and interpretive character of assessment and potentially devalue teachers' professional expertise.

## Study III: Hello GPT! Goodbye home examination?

An exploratory study of AI chatbots' impact on university teachers' assessment practices

Study III examines the disruptive role of generative AI (GAI) in higher education, with particular attention to how ChatGPT mediates university teachers' assessment practices. The launch of ChatGPT in November 2022 has sparked discussions globally about the influence of GAI chatbots on educational practices within higher education institutions. This study examines teachers' perceptions of ChatGPT's performance in responding to home examination prompts within undergraduate contexts and explores the mediateing role of GAI chatbots on teachers' assessment practices. Involving 24 university teachers from humanities and social sciences faculties, four Turing Test-inspired experiments were conducted where participants were asked to blindly

assess responses to home examination questions, both from students and ChatGPT (GPT 3.5, including updates until March 2023).

ChatGPT-generated text achieved a passing rate ranging from 37.5% to 85.7%. The participants highlighted several strengths in the ChatGPT responses such as competence in language, precision, consistency and creativity. Identified weaknesses included poor argumentation, limited referencing of course literature, and a lack of alignment with the course content. However, such flaws could also be found in student texts. Several aspects of the texts led the participants to 'flag' and suspect that the text was not written by a human (14-23%), such as the lack of personal voice, clear argumentation, and connections to course content. More specifically, the participants reported that the ChatGPT-generated texts often contained nonsensical statements, lacked personal opinions and emotional expression, included factual inaccuracies, and made odd use of synonyms or translations of English terms into Swedish. They also noted that the texts tended to repeat the examination question prompt, displayed stylistic features commonly associated with generative AI chatbot outputs from previous experience, and were frequently inconclusive in their arguments.

Additionally, the findings of the study shows that the participants demonstrated a more critical approach to grading student texts, with lower scores assigned to responses characterised by repetition, limited length, and insufficient use of course literature. The participants seemed to raise their standards for awarding high and passing grades, as in rare circumstances they awarded an A and in total failed more than 9% of student responses, which had previously received passing grades. More specifically, the participants assigned low scores to student texts due to having repetitive statements, incoherence in content and meaning of the text, mixing accurate and inaccurate claims, and having short in length texts. The participants also assigned lower scores due to mistakes with referencing, including missing citations and limited engagement with relevant literature beyond the course materials. Finally, the participants suspected that the student texts were GAI-chatbot-written when the texts seemed to include inaccuracies, perceived as non-human, or assumed that the text was 'too good to be written by a student'.

Drawing on postphenomenology, the study argues that ChatGPT mediates assessment practices by altering teachers' perceptions of student work. Participants reported being constantly aware of the possibility that a text they were assessing could be AI-written, which shaped their evaluations. This study shows that the presence of GAI chatbots may prompt teachers to ask "who has written the text?" and thereby question students' authorship, potentially reinforcing mistrust at the core of the teacher–student relationship.

## Study IV: Navigating Uncertainty

University teachers' experiences and perceptions of generative AI in teaching and learning

Study IV explores how generative AI (GAI) mediates teachers' practices and reconfigures their professional roles, focusing on university teachers' experiences and perceptions of GAI in higher education. This study draws on postphenomenology and technology mediation theory (Verbeek 2005, 2011; Rosenberger & Verbeek, 2015), which emphasise how technologies are not neutral instruments but artefacts that shape perception, action, and morality, and allow GAI to be understood as multistable where its meaning and purpose are negotiated within educational contexts. The research was conducted at a large Swedish university with 24 teachers from sociology, philosophy, law, and education, disciplines where long-format essays constitute central elements of assessment. In workshops during spring 2023, participants first encountered GAI-generated outputs from their own course assignments and then took part in focus group interviews. The material was analysed thematically following Graneheim and Lundman (2004), producing six key themes: three concerning teachers' experiences of the emergence of GAI (alarming and overwhelming situation; a state of vulnerability; and emergence ethical challenges) and three concerning teachers' perceptions of the potential transformation in teaching and learning practices (rethinking assessment; re-evaluating teaching priorities; and student learning at risk).

Teachers described the emergence of GAI artefacts as sudden and overwhelming, while a "hype-driven" media discourse portrayed students as likely cheaters, intensifying a sense of urgency and crisis in higher education. Several teachers expressed uncertainty about the capabilities and limitations of GAI and anxiety about being technologically surpassed by their students, highlighting that less experienced academics were considered especially disadvantaged, lacking familiarity with course-specific assessment practices that might help them identify authentic student work. Such uncertainties often left participants feeling vulnerable and exposed to mistakes in teaching and assessment. The emergence of ethical challenges was another central theme. Teachers voiced concern about ensuring fairness in assessment, avoiding wrongful suspicion of students, and preserving trust within the teacher–student relationship. Unequal access to AI tools, due to the distinction between free and paid versions, was regarded as a source of inequity among students. Participants also raised issues of bias and inaccuracy in GAI outputs, echoing findings from Kasneci et al. (2023), and highlighted the danger that students might accept AI outputs uncritically as reliable knowledge.

Alongside these experiences, teachers reflected on possible transformations in teaching and learning, with assessment being at the centre of these

discussions. Several participants suggested abandoning home examinations in favour of oral or in-class formats, while others proposed incremental submissions or diversified assessment modes to make AI reliance less practical. Some emphasised transparency, arguing that students should declare their use of AI or sign statements of non-use, equating undisclosed use with plagiarism. Re-evaluating teaching priorities was also seen as essential, where the participants underlined the need to foster critical thinking, evaluative judgement, and ethical engagement with technology, aligning with calls in the literature for supporting students' capacity to navigate GAI critically (Bearman et al., 2024). Several participants noted possible pedagogical uses, such as designing exercises where students critique AI-generated responses against course materials, thereby exposing inaccuracies and encouraging deeper engagement with subject content. At the same time, the participants stressed their responsibility as teachers to guide students in distinguishing between legitimate assistance and plagiarism, underscoring academic integrity as a central concern. In addition, several concerns in relation to student learning and the potential of students losing the ability to respond independently of GAI, fearing that GAI could erode core academic skills such as writing, reasoning, and doing research. They also raised concerns about the desocialisation of learning, fearing that students might turn to AI rather than peers, thereby weakening collaborative knowledge construction.

The findings suggest that teachers perceive the emergence of GAI as both disruptive and potentially transformative. They experience heightened feelings of vulnerability, inadequacy, and ethical tension, while also recognising opportunities to re-examine fundamental assumptions about knowledge, assessment, and the purpose of teaching. These experiences resonate with broader accounts of uncertainty as a defining feature of the current AI landscape (Barnett, 2012; Bonnet & Glazier, 2024a). In conclusion, the study contributes to the growing body of research on GAI in higher education by foregrounding teachers' voices at a moment of transition. It demonstrates how the emergence of GAI unsettles confidence and challenges established norms, while also creating openings for pedagogical renewal. The findings of the study show that the participants were compelled to reconsider assessment formats, teaching priorities, and their own roles in fostering critical and ethical engagement with technology among their students. Yet their limited preparedness and insufficient support leave them vulnerable. The findings underline the importance of strengthening AI literacy and creating collaborative spaces for critical exploration if higher education is to navigate the uncertainties of GAI meaningfully.

# Discussion

This chapter brings together the contributions of the studies included in the thesis, discussing how they respond to the thesis' research questions. Additionally, this chapter offers an overview of the thesis' contribution to research, presents the limitations and delimitations of the thesis and concludes with the implications for practice and research.

## Teachers experiencing the emergence of AI

The first research question (*How do university teachers experience the emergence of AI in the context of their practices?*) focuses on teachers' experiences of AI's emergence, foregrounding how AI technologies are encountered and interpreted in relation to their practices (*experiential level*). To address this research question, I examined how university teachers experience and interpret the emergence of automated grading systems (AGS) and generative AI (GAI) chatbots in the context of their practices, drawing on the findings in Studies II, III and IV. In doing so, it was necessary to account for two important contextual considerations: (a) differences between the technological artefacts that are being examined, and (b) differences between the participants, particularly between those involved in Study II and those in Studies III and IV.

With regard to the differences between the technological artefacts, both AGS and GAI chatbots are considered as emerging technologies in my research but have several distinctive characteristics that relate to the ways the participants experienced and interpreted their emergence in their practices. First, the AGS studied in my project refer to 'in-house' developed tools, tailored to the teachers' needs for a specific course, while GAI chatbots refer to publicly available 'general purpose' commercial products designed by private companies. More specifically, the AGS in focus in Study II are tools developed by the teachers themselves and designed to meet a need and/or improve their current administrative and assessment practices. In contrast to the AGS, GAI chatbots were not designed within and for higher education by teachers themselves, by colleagues, or by higher education technology providers to

meet an educational need. Instead, they landed as an external force from private companies aiming to gain popularity and profit and not particularly to serve educational benefits. Second, AGS can be considered as an example of an AIED technology, falling under the "teacher-focused" for "assessment and evaluation" group of technologies according to (Holmes & Tuomi, 2022) and Zawacki-Richter et al. 's (2019) classifications of AIED systems. Under these terms, AGS is a technology that was designed with the purpose to fulfil particular educational goals. Unlike AGS, GAI chatbots could not be considered as AIED systems, since they were not designed for educational contexts and purposes (Cerratto Pargman et al., 2024), and are rather promoted and positioned in the market as conversational agents, capable of responding to any type of user prompt and can be used by anyone for any reason. GAI chatbots, therefore, cannot be classified as educational technologies per se but rather as general-purpose AI technologies.

Differences between participants in Study II and in Studies III and IV concerned both disciplinary background and the timing and conditions under which technological artefacts emerged in relation to their practices. Although all participants were university teachers, those in Study II were computer scientists, whereas participants in Studies III and IV were affiliated with departments of law, philosophy, sociology, and education. In Study II, teachers were prepared for the testing and deployment of the AGS within their own courses. Additionally, the study was situated in a computer science context in which emerging technologies constitute an integral part of both teaching and research. Accordingly, participants possessed high levels of technical expertise in areas such as programming, data mining, machine learning, and data science. Such a technical expertise seems to be related to a) their familiarity in terms of emerging technologies being an integral part of their area of teaching and research, as well as b) their agency in piloting emerging technologies in their practices, in terms of their capacity to understand the technicity of such technologies and autonomy in tweaking the technologies to fit and meet their needs. These elements seem to be related to a high degree of familiarity with the AGS in question as well as to a sense of confidence, curiosity, and willingness to experiment with AGS in practice. By contrast, the participants in Studies III and IV encountered the emergence of GAI chatbots from a markedly different position. Teachers in philosophy, sociology, law, and education reported limited familiarity with the technicity of GAI systems and, in many cases, stated that they were unaware of what such artefacts could achieve in the context of higher education. Unlike the participants in Study II, they did not participate in the design or deployment of the technology, nor did they experience its emergence as a gradual or internally driven process, but rather as an unfamiliar and externally driven technology.

Considering these differences and drawing on the findings from Studies II–IV, the findings indicate a clear contrast in how the two technologies were experienced. Whereas the teachers experiencing the emergence of AGS did so

with a high degree of familiarity and agency, allowing them to experiment with such artefacts in their practices, the emergence of GAI chatbots was experienced as more abrupt and uncertain, shaped by unfamiliarity and reduced perceived agency. More specifically, in contrast with the case of AGS, where the participants of Study II experienced its emergence with a certain optimism, built on the promises of technology and expectations from their crafted artefacts, in the case of GAI chatbots, the participants of Studies III and IV described how their emergence fuelled a sense of 'fear to the unknown'.

Taking a closer look to the findings of the studies, in the case of AGS, such preparedness of the participants was to a great extent infused by expectations and imagined outcomes (*promises*), such as increasing their work efficiency and improving the quality of their assessment. However, AGS, when practised and experienced by the teachers, introduced several disruptions and unwanted consequences (*breakages*) to their practices that were not anticipated. Although the emergence of AGS was expected by the participants, when they piloted and deployed their artefacts in their practices, they experienced AGS mediations as ambivalent – both as promising and disruptive. Additionally, in the case of GAI chatbots, teachers appeared largely unprepared and alarmed, encountering the emergence of a highly hyped technology with discomfort, particularly in relation to its perceived capabilities. This lack of familiarity contributed to heightened *uncertainty* about how to respond to and cope with such technologies in higher education contexts, thereby producing a sense of *vulnerability* marked by reduced perceived agency. This sense of vulnerability was reflected in participants' concerns about their own professional judgement, including the risk of adopting overly critical or punitive stances toward students, such as falsely suspecting them of using GAI chatbots in written assignments.

According to Bearman et al. (2022), "how universities respond to AI depends not only on what AI is but also on what it is *understood* to be" (p. 370). In other words, responses to AI are shaped by socially and pedagogically co-constructed meanings, rather than by the technical properties of AI technologies alone. From this perspective, university teachers play a critical role, as their interpretations of AI shape how such technologies are accepted, resisted, or adapted within teaching practices. Consequently, exploring how university teachers perceive the emergence of AI technologies can provide important insights into how these key stakeholders respond to AI in higher education contexts. Moreover, examining technologies at the point of their emergence makes visible the reasons, ways, processes, and conditions through which technological artefacts become embedded in higher education practices – or fail to do so.

I address this first research question by drawing on technology mediation theory and adopting a relational approach (Verbeek, 2005; 2011). This approach foregrounds the ways that educational practices, experiences, and responsibilities emerge through ongoing human-technology-world relations,

thereby opening up *opportunities*[16] to further explore the entanglements of teachers, practices, technologies, and ethics within AI mediations. Additionally, by not perceiving AGS and GAI chatbots as mere tools, but rather as influencing teachers' 'existence and experience' (Rosenberger & Verbeek, 2015) in the context of their practices, I examined how these artefacts shape and are shaped by teachers' perceptions when they emerged into their practices. While AGS was perceived as a promising solution to existing issues related to time and work management, as well as the quality of assessment practices, GAI chatbots were perceived as an external force invading teachers' practices. In both cases, however, the emergence of AI was understood as introducing wanted or unwanted disruptions and challenges into higher education practices.

Across the three studies (II, III, IV), the findings point to different ways in which university teachers experience the emergence of AGS and GAI chatbots within higher education practices. Despite these differences, two common themes were identified across both cases. Both technologies were perceived as introducing certain *changes* and *ethical issues* in established practices, particularly in relation to teachers' roles and responsibilities and their relationships with students. These changes – whether experienced as forced or voluntary – were perceived as disruptive to how the teachers understand and enact their practices. As shown in Studies II-IV, these disruptions are not inherently uniform in how they are valued; rather, they are experienced differently depending on the artefact in question, its purpose and design, and stakeholders' levels of AI literacy[17] and readiness[18]. Such disruptions were perceived as expected or unexpected, desired or undesired, and welcomed or resisted, bearing either promises or threats when introduced into higher education practices. These findings echo a broader pattern in previous research, in which AI in higher education is variously framed as a constructive innovation or as a disruptive and potentially problematic force, but mainly as introducing disruptions to established practices (Bearman et al., 2022). In addition, both AGS and GAI chatbots raised several ethical issues, including risks, challenges, and concerns in relation to AI-mediated higher education practices.

---

[16] The notion of *opportunities* here, aligns with the notion of *freedom* – 'spaces and possibilities that enable and constrain certain actions and perceptions' (Aagaard, 2017, p. 527) – that Verbeek (2011) and Aagaard refer to in relation to the influence of technologies in human practices.

[17]According to Long & Magerko (2020), AI literacy is defined as "a set of competencies that enables individuals to critically evaluate AI technologies; communicate and collaborate effectively with AI; and use AIas a tool online, at home, and in the workplace" (p. 598)

[18] According to Luckin et al. (2022) "The concept of AI Readiness is a way to describe the transition that those working in education and their students need to make from not understanding what AI is and what AI can do, to being able to understand, in non-technical terms, what AI is capable of achieving." (p. 1)

The second research question addresses these changes (*perceptual level*), while the third research question focuses on the ethical issues introduced by AGS and GAI chatbots from teachers' perspectives (*ethical level*). Both are examined in detail in the following sections.

## AI reconfiguring teachers' perceptions of higher education practices

While the previous section focused on the experiential encounter of university teachers with the emergence of AI, the second research question (*How does AI reconfigure university teachers' perceptions of their practices?*) focuses on how these encounters lead to potential reconfigurations of teachers' perceptions of higher education practices (*perceptual level*). To address this research question, I explored how AGS and GAI chatbots reconfigure university teachers' perceptions of their roles, responsibilities and relationships with students in the context of their practices.

Drawing on the findings in Studies II, III and IV, this section builds on the previous analysis, which showed that teachers perceived AI as introducing certain changes and disruptions in higher education practices. Here, these changes are conceptualised as *reconfigurations* of AI-mediated practices relating to (i) new forms of labour; (ii) instrumental learning incentives; (iii) educational judgement; and (iv) student-teacher relationship.

### AI and new forms of labour

AI reconfigures teachers' perceptions of their work by introducing new, often less visible, forms of labour into their practices. In Study II, although AGS were initially expected to enhance teachers' efficiency by streamlining assessment and reducing grading time, participants reported that their use, in practice, produced additional frictions around the preparation and continuous revision of assessment material. AGS-mediated assessment practices required teachers to design assessment questions that conformed to the system's constraints or specificities, a process perceived as disproportionately time-consuming and effort-intensive. In addition, the structural inflexibility of AGS restricted teachers' capacity to revise or update assessment material on a yearly basis, even when they identified a pedagogical need for such changes. Such rigidity impeded their ability to ensure responsiveness and continuous improvement in teaching, while attempts to scale the use of AGS across multiple courses were also described as labour-intensive.

AGS mediations within assessment practice seemed to also shift the locus of effort from grading towards design and maintenance, thereby *redistributing rather than reducing teachers' workload* and, in turn, adding complications to their pedagogical practice. Similar conclusions are drawn by Selwyn et al. (2023), who argue that the introduction of automation in education practices calls teachers to engage in new forms of work, such as "fitting their actions around what the machine is capable of recognising" (p. 17) and developing so-called "parseable" pedagogies that render teaching practices legible to automated systems. Such analysis also resonates with Bainbridge's (1983) notion of the ironies of automation, which highlights how technological systems often extend rather than eliminate the problems in the tasks they are designed for. As with Bainbridge's (1983) mechanical automation requiring "human operators" to take on additional tasks, such as monitoring, maintenance, and intervention in cases of failure, AGS seemed to demand teachers to be attentive to potential mistakes, limitations and biases of the systems adding a new set of tasks to follow. Thus AGS-mediated assessment can be understood as producing not a reduction of labour, but a reconfiguration of academic work.

Similar to AGS, the use of GAI chatbots by teachers within their practices also appears as labour-intensive, albeit in less direct ways. Recent research suggests that teachers' engagement with GAI involves substantial hidden work that is often framed as routine "checking" or "processing" but in practice requires considerable time and effort (Selwyn et al., 2025). This work includes exercising evaluative judgement over AI-generated outputs, deciding when such outputs are no longer improving, and manually editing or reworking materials to meet pedagogical aims. Finally, both AGS and GAI chatbots' use can be understood as reconfiguring academic work by redistributing effort toward design, maintenance, oversight, interpretation, and intervention. In this sense, AI-mediation in higher education practices do not eliminate labour but transform its nature, rendering some forms of work less visible while intensifying demands on teachers' professional judgement and attentiveness.

## AI and instrumental learning incentives

Teachers perceived AI as reshaping students' engagement with course content, encouraging a more instrumental approach to learning focused on task completion rather than deep understanding. In Study II, when AGS were used to mediate assessment and feedback, participants reported that students increasingly oriented their efforts towards meeting the system's functional requirements, such as producing code that compiled, ran, or satisfied the predefined test criteria. AGS-mediated assessment practices appeared to encourage students to prioritise task completion over engagement with underlying programming concepts, often progressing through assignments

with limited attention to structure, design, or problem-solving processes. In addition, participants described how system-generated feedback was frequently approached by students as a series of technical issues to be resolved rather than as an opportunity for reflection and learning. Such feedback iterations were perceived as narrowing students' focus to what needed to be corrected in order to advance, rather than supporting a more comprehensive understanding of mistakes. Furthermore, participants described how some students treated features such as unit tests as 'crutches' or procedural aids for completing assignments, without necessarily engaging with the underlying concepts being assessed.

Such AGS mediations within assessment practice seemed to discourage experimentation, reflection, and thinking beyond the immediate task, thereby *reinforcing instrumental learning incentives* and, in turn, reconfiguring how learning was framed and enacted within assessment practice. This view resonates with the recently growing field of critical studies of AI in education (Holmes et al., 2025) which probematises such a transactional character of teaching and learning activities, as it is often promoted in research and policy on AI integration in education. Similar dynamics have been identified in studies examining the use of GAI chatbots by students for feedback and solution-seeking. When such tools are primarily used to obtain answers or ready-made solutions, they may undermine reflective learning processes and motivation for learning (Fan et al., 2024), while reducing students' engagement with content at a deeper conceptual level (Stadler et al., 2024). Moreover, empirical research indicates that students' motivations for using AI in higher education are frequently linked to time-saving, time management, and performance enhancement (Morell-Mengual et al., 2025), further facilitating and reinforcing instrumental views of learning within AI-mediated educational practices.

## AI and educational judgement

AI unsettles teachers' perceptions of the purpose and function of their existing practices, prompting them to rethink their teaching priorities. In Study IV, the participants perceived that the emergence of GAI chatbots required them to *rethink their assessments* and *re-evaluate their teaching priorities*. More specifically, they highlighted that the presence of GAI chatbots unsettled established understandings of assessment by calling into question its role both as a mechanism of validation and as a process of learning. This was perceived by the participants as requiring a reconsideration of what assessment is for and how it should be designed. Such a rethinking of assessment purposes and re-evaluation of priorities can be understood not only as a technical or procedural fix that the teachers perceived as necessary, but as one that directly engages teachers' professional judge-

ment. Such judgement in higher education is characterised by context sensitivity, responsiveness, and the ability to balance formal criteria with situated understanding and reflexes (Macfarlane, 2003).

Additionally, such reconfiguration of how teachers perceived their roles and tasks introduced frictions into their practices, as they were compelled to navigate tensions between long-standing assessment logics and the new pedagogical demands brought about by AI mediation. Participants emphasised an increasing need to prioritise students' development of critical, reflective, and evaluative skills for studying with – and eventually working alongside – GAI chatbots. In this sense, GAI chatbot mediations not only reshaped teachers' perceptions of their assessment-related tasks but also reframed their broader responsibilities within teaching and learning. This aligns with emerging scholarship (e.g., Holmes et al., 2025; Holmes & Tuomi, 2022; Selwyn et al., 2025) that frames AI in education as a catalyst for renewed attention to pedagogical judgement, rather than its obsolescence, highlighting the continued centrality of educators' professional decision-making in AI-mediated learning environments.

## AI and student-teacher relationship

AI contributes to the amplification of *mistrust* in student-teacher relations in the context of higher education practices. In Study III, participants reported being constantly aware of the potential presence of AI-generated content in student work, which led them to adopt a more critical stance toward students' texts. This shift often translated into raising the threshold for what was considered a passing grade and adopting less tolerance for flaws in language and structure, thereby reconfiguring their evaluative practices. Findings from both the experiment and participants' subsequent reflections further revealed a sense of 'policing' tension, as teachers felt compelled to scrutinise and flag texts they suspected to be AI-generated. These experiences suggest that GAI chatbots not only altered how teachers assessed student texts but also reconfigured their perceptions of *responsibility in maintaining academic integrity*. Echoing the findings of Study III in relation to GAI and assessment of student work, further research has examined such "erosion of trust" in student-teacher relationships (Luo (Jess), 2025) and policing mindset (Giray et al., 2025), driven by concerns about cheating and uncertainty over authorship.

Finally, AI can contribute to desocialisation of assessment practice. Drawing on Study II, participants raised concerns about AGS-mediated assessment practices leaving less space for student-teacher interaction and contributing to a loss of personal connection. From a technology mediation perspective (Verbeek, 2005, 2011), this can be understood as AGS actively mediating relational dynamics, as they do not simply "support" assessment

but reconfigure what teachers and students attend to, how responsibility is distributed, and which forms of interaction become more or less likely.

AI artefacts mediate university teaching practices by shaping how teachers perceive, enact, and make sense of their professional roles. In line with post-phenomenological accounts of technological mediation (Verbeek, 2005; 2011), such technologies can be understood as co-constitutive of professional roles and relations, which actively shape teachers' experiences and practices by mediating how they perceive, enact, and make sense of their roles, how they practice assessment, and how they understand their responsibilities. Drawing on Studies II-IV, AGS and GAI chatbots do not simply support or disrupt existing practices but actively co-shape them by redistributing attention, effort, and responsibility. Across the studies, AGS mediations were experienced as reorienting teachers' labour towards design, maintenance, and oversight, while simultaneously foregrounding instrumental learning incentives within assessment practices. At the same time, GAI chatbot mediations prompted shifts in educational judgement, by mediating how assessment was understood as both a process of validation and a site of learning, and student-teacher relationship, by amplifying mistrust.

Rather than being read as a list of what AI can or cannot do when practised in higher education, the reconfigurations outlined in this section should be understood as analytical pathways for making sense of how AI mediations shape university teachers' perceptions of their roles, responsibilities, and relationships in the context of higher education practices. From a technology mediation perspective, these pathways do not represent external impacts or linear consequences of AI adoption, but instead point to how the emergence of AI interacts with, and gives new form to, pre-existing tensions within higher education practices. Issues such as workload, assessment purposes, learning orientations, professional judgement, and trust were not introduced by AI as such; rather, they were mirrored, amplified, or distorted (Vallor, 2024) through AI-mediated practices. In this sense, the meaning and significance of AGS and GAI chatbots are not inherent to the technologies themselves but are co-constituted through how teachers engage with them in situated practices. These engagements are shaped and influenced by broader sociotechnical conditions and environments, including public discourses around AI, national imaginaries of automation and efficiency, institutional values and strategies, as well as the everyday needs, concerns, and challenges of academic work. Attending to these mediations foregrounds AI as a relational technology that participates in shaping how teaching and learning contexts, purposes and conditions are perceived, enacted, and valued. Building on this understanding, the following section turns more explicitly to the ethical issues raised within AI mediations in higher education practices.

## Ethical issues emerging in AI-mediated practices

Building on the previous sections focusing on the experiential and perceptual levels of university teachers' AI-mediated higher education practices, the third and final research question (*What ethical issues emerge when AI is introduced into higher education practices?*) focuses on the ethical issues that arise in the context of such mediations (*ethical level*). To address this research question, I explored ethical issues, including risks, challenges, and concerns that take shape in the context of AI-mediated practices for teachers and students, focusing on AGS and GAI chatbots. All four studies conducted for this thesis address this research question; therefore, in this section, I present the main ethical issues identified in Studies I-IV, which relate to *data ownership*, *bias perpetuation*, *displacement of ethical responsibility in assessment*, *academic integrity and authorship*, *fair treatment of students* and *access to AI*. Rather than constituting an exhaustive or prescriptive list of challenges to be addressed, these issues are presented as key areas of ethical problematisation that emerged across the studies conducted in this thesis.

### Data ownership

Issues of data ownership constitute a central ethical concern in the context of AGS, particularly in relation to *students' autonomy and agency*. As data-driven systems increasingly shape assessment practices, questions arise regarding who controls student-generated data, under what conditions it is used, and how consent is obtained. These questions are especially pressing where students may be unaware that AGS mediate assessment processes, and when students may have limited opportunity or capacity to give informed consent to the use of their data.

More specifically, Study I points out that AGS are often built and trained on student data, such as grades, assignments, and examination responses. This raises important questions about informed consent, as students could often be in a position to not fully understand or negotiate how their data will be used, for how long, and for what purposes. Existing power relations between students and educational institutions further complicate this issue, as consent may be given under implicit pressure rather than as a free choice. Such observation also aligns with Holmes et al. (2025) noting that "students are rarely consulted or included in decisions about the impact of AI and how AI systems shape their opportunities and choices" (p. 19). Beyond questions of consent, the use of students' work to build AGS also raises issues of *data ownership* and *intellectual property*. As discussed in earlier research (Bearman et al., 2020; Cerratto Pargman et al., 2023), students' contributions can be understood as a form of 'invisible labour' that under-

pins the functioning of these systems while blurring the boundaries between individual and institutional property. Accordingly, ensuring students' autonomy and agency requires not only making consent meaningful and revocable but also designing AGS in ways that respect students' rights to withdraw their data and actively participate in decisions about the technologies that directly affect them.

From a relational ethics perspective, these concerns cannot be reduced to questions of data protection or compliance alone, but need to be understood as emerging through the mediated relationships between students, institutions, and AI artefacts. AGS do not merely process student data but actively shape how agency, ownership, and responsibility are distributed within assessment practices. Ethical engagement with data ownership, therefore, requires attention to how these technologies reconfigure relationships and power dynamics, rather than treating consent as a one-off procedural safeguard. Finally, reflecting on GAI chatbots as an emerging technology that could be related to assessment and feedback practices (Flodén, 2025; Henderson et al., 2025; Xia et al., 2024), similar ethical questions around data ownership and consent also arise in relation to sharing students' work with GAI chatbots, particularly where student-generated content may be retained, reused, or repurposed within opaque data infrastructures beyond higher education stakeholders' direct control.

## Bias perpetuation

Ensuring fairness in AGS-mediated assessment has been framed as a major ethical challenge for university teachers both in previous research (Holmes & Porayska-Pomsta, 2022; Kizilcec & Lee, 2022) and in the empirical findings of this thesis. In Study II, the participants described balancing the biases of AGS as a recurring tension or "breakage" introduced into assessment practice. While AGS were initially envisioned to enhance fairness by ensuring consistency and reducing human subjectivity, participants highlighted that these systems often replicate and even amplify existing biases embedded in historical data, training strategies, and individual grading styles. The reliance on historical data further raised concerns that AGS-mediated assessment may not necessarily be more fair or human-bias-free than human grading, particularly when a single teacher's grading style or a limited set of annotated responses informs the system's "ground truth".

These concerns resonate with Study I, demonstrating that AGS are typically trained on previously graded student responses and therefore inherit the values, assumptions, and limitations of past assessment practices. Hence, biases that may have remained implicit or unnoticed in human grading risk becoming stabilised and scaled through automation. This underscored that *the pursuit of fairness cannot be reduced to the removal of hu-*

*man subjectivity through automation*. Instead, as both Studies I and II illustrate, teachers are compelled to navigate a complex ethical terrain in which the *promise of algorithmic objectivity coexists with the practical risk of reinforcing bias*.

These findings highlight that questions of fairness and bias emerge through the ongoing relationships between teachers, students, and AI systems in the context of higher education practices, rather than residing solely within the technology itself. AGS mediate how past judgements are carried forward into present assessment practices, thereby shaping whose knowledge, performance, and learning trajectories are recognised or marginalised. Ethical concerns around bias thus arise not only from technical design choices but from how responsibility, judgement, and authority are redistributed across human and AI artefacts.

## Displacement of ethical responsibility in assessment

One ethical issue emerging across the studies concerns the displacement of ethical responsibility in assessment practices when assessment is mediated by AI. In Studies I and II, AGS were primarily framed in relation to efficiency, scalability, and the acceleration of feedback, foregrounding procedural aspects of assessment over its pedagogical dimensions. More specifically, as Study II identified, managing teachers' workload, in terms of reducing their workload and saving time, was one of the main drives for developing and using AGS. Additionally, the findings of Study II show that another driver of AGS development and use by teachers is to improve assessment, in terms of fairness, consistency and reducing human bias. According to the participants in Study II, removing the human fallible subjectivity in the grading process and trusting a non-human and perceived as less-likely-to-make-mistakes system, was considered more safe and fair. While such framings seem to respond to the increased workload and demands on university teachers they also risk hollowing out assessment as a pedagogical practice grounded in professional judgement, dialogue, and accountability.

A central pillar of such perceptions seems to be the persistent appeal to objectivity as an ethical ideal in assessment. Participants in Study II contrasted human subjectivity – associated with fatigue, inconsistency, or bias – with the expected neutrality and reliability of AGS. However, as demonstrated in Studies I and II, this aspiration rests on an imagined form of mechanical objectivity that obscures the value-laden and interpretive nature of teaching and assessment practices (Harland & Pickering, 2011; Macfarlane, 2003). In line with Porter's (2020) conceptualisation of mechanical objectivity as knowledge "based completely on explicit rules" (p. 7) – an ideal that is theoretically unattainable – AGS, as discussed above,

do not eliminate bias but rather reconfigure it through historical data, design choices, and algorithmic classifications, while simultaneously rendering these judgements less visible and less contestable. In this sense, the *pursuit of fairness through automation with AI risks displacing ethical responsibility from situated professional judgement to technical systems whose biases are harder to identify and scrutinise*.

Echoing concerns raised by Selwyn et al. (2023) regarding the automation of educational decision-making, such a shift could contribute to a broader devaluation of teachers' educational judgement. Even when positioned as decision-support tools, AGS risk reshaping what counts as legitimate assessment work, privileging what is easily measurable, standardisable, and automatable. Over time, this may normalise assessment practices that bypass teachers' interpretive engagement with student work, encouraging overreliance on automated outputs and narrowing the scope of what is assessed. Study I also highlights that AGS mediations that minimise or bypass teachers' engagement in assessment might introduce notable frictions. These frictions primarily concern the risk of teachers overlooking critical signs of students' performance, which can result in inadequate feedback and diminished support for learners. Furthermore, *overreliance on AGS judgments may reduce opportunities for teachers to exercise and refine their own educational judgment*, potentially weakening their professional expertise over time. Such developments exemplify a paradox of AI-driven automation, where although AGS are introduced to support teachers, they simultaneously risk undermining the moral, relational, and pedagogical dimensions of assessment that require human expertise, accountability, and care.

The displacement of ethical responsibility in assessment can be understood through technology mediation theory (Verbeek, 2005; 2011) as emerging through how AGS reconfigure relationships between teachers, students, and assessment practices. By privileging procedural efficiency, standardisation, and the ideal of mechanical objectivity, AGS mediate what forms of judgement, care, and accountability become visible and valued in assessment. The ethical concern here, then, is not automation per se, but how responsibility and moral engagement are displaced from situated pedagogical relationships to AI artefacts. The displacement of ethical responsibility in assessment is therefore not merely a matter of efficiency gains or technological limitations, but an ethical issue rooted in how assessment is re-imagined in the context of AI-mediated higher education practices. When assessment is reframed as a procedural task optimised through automation, its role as a socially situated, interpretive, and relational practice is weakened, raising fundamental questions about accountability, trust, and the purposes of assessment in higher education.

## Academic integrity and authorship

Issues around academic integrity and authorship were of high concern in relation to GAI chatbots in higher education. More specifically, participants in Studies III and IV highlighted two major areas of concern: first, regarding what should be considered *acceptable* or *unacceptable* use of AI in student submissions for formal examinations, and second, regarding what should constitute *ownership* and *authorship* in student texts produced with AI assistance.

In both studies, several participants equated students' use of GAI chatbots with cheating, framing it as a new form of plagiarism in assignment writing. This aligns with a significant volume of literature that conceptualises student use of GAI for examination tasks as a form of academic dishonesty, particularly when such use is not disclosed or violates existing assessment regulations (e.g. Cotton et al., 2024; Perkins & Roe, 2024). At the same time, some participants pointed out that GAI chatbots are widely accessible to the public, including students, and may therefore be compared to other technological artefacts – such as calculators, spelling and grammar checkers, or internet-based sources – that are not generally regarded as illegitimate tools in academic work. Recent studies have similarly argued that GAI challenges binary distinctions between cheating and legitimate assistance, suggesting that its ethical status depends on context, transparency, and pedagogical intent rather than on the technology itself (e.g. Corbin et al., 2025; Luo & Dawson, 2025). This tension was perceived by participants as a call for university teachers, and higher education institutions more broadly, to revisit and clarify existing rules and regulations concerning what constitutes legitimate and illegitimate uses of AI by students. Reflecting these concerns, the Swedish Higher Education Authority (UKÄ) has reported a sharp increase in the number of disciplinary cases related to 'AI plagiarism' –unauthorised use of AI tools – in higher education in 2023 and 2024 (Kyrk et al., 2024). Hence, it can be argued that the emergence of GAI and its introduction to higher education has unsettled existing regulatory frameworks and intensified uncertainty around academic integrity in assessment practices.

Beyond questions of rule compliance, participants also raised deeper concerns regarding authorship and ownership, questioning whether students can be said to fully 'own' a text that has been co-authored, in whole or in part, by a GAI chatbot. This problematises traditional understandings of authorship that underpin academic integrity, originality, and intellectual property in higher education. As Macfarlane (2003) argues, academic integrity constitutes "a cornerstone of academic life" (p. 29), essential for maintaining trust and credibility within academic communities. Importantly, Macfarlane (2003) emphasises that integrity extends beyond the avoidance of plagiarism or misconduct, involving a broader commitment

to truth-seeking and intellectual honesty. From this perspective, the ethical challenges posed from GAI chatbots' emergence in higher education practices are not limited to whether AI use can be detected or sanctioned, but concern how originality, responsibility, authorship, and scholarly accountability are reconfigured when writing practices are increasingly mediated by AI.

## Fair treatment of students

Another ethical issue in relation to GAI chatbots in Studies III and IV concerns the risk of unfair treatment of student work by teachers. The findings of Study III indicate that the emergence of GAI chatbots may amplify teachers' *criticality* when assessing student texts, leading to heightened *suspicion* that students have engaged in "cheating" through AI use. In such cases, students who may have used GAI chatbots to draft their assignments – or who are merely suspected of doing so – risk being labelled as "cheaters", with teachers potentially flagging or penalising their texts in the absence of clear evidence. Drawing on the findings from both studies, several concerns were raised in relation to such dynamic risking to undermine fairness, as students may lose points or be judged more harshly because teachers, unconsciously or deliberately, raise standards when assessing student work. These dynamics echo recent research showing that GAI can contribute to undermining trust in teacher–student relationships, particularly where suspicion of AI use becomes a dominant interpretive frame (Luo (Jess), 2025). Previous research further cautions that policing-oriented approaches to AI in higher education practices can foster environments of distrust and anxiety and disproportionately penalise students without enhancing genuine academic integrity (Giray et al., 2025).

The ethical challenge here lies in balancing the responsibility to uphold academic integrity with the equally important responsibility to avoid unfair suspicion, over-criticality, or punitive treatment of students. From a relational ethics perspective, fairness in assessment depends not only on preventing misconduct but also on sustaining trust and care in educational judgements – values that risk being compromised when suspicion becomes a dominant lens for interpreting student work.

## Access to AI

Drawing on Study IV, another ethical concern associated with the emergence of GAI chatbots relates to issues of equity and access. More specifically, participants raised concerns about students' *unequal access* to GAI chatbots, particularly the divide between those who can afford premium subscription plans, offering more advanced features, and those limited to free versions with restricted functionality and access to less sophisticated

models. This uneven access was perceived as potentially widening existing inequities between students, granting some a technological advantage in their use of GAI chatbots for learning-related purposes. The ethical concern, therefore, lies in the risk that the introduction of GAI chatbots into higher education practices may exacerbate social and economic inequalities, shifting advantages toward students with greater financial resources.

Finally, it is important to reflect that GAI chatbots are developed and operated by private companies that offer market-based access structures, risking amplifying inequalities, as students' opportunities to access such AI artefacts can vary significantly. This also raises questions about the reliance of higher education on commercially driven technologies that may not be open source, may operate through opaque data practices, and may extract value from student interactions for commercial purposes. In this context, issues of access cannot be separated from broader questions of institutional responsibility, particularly regarding the integration of GAI chatbots in higher education practices that shape students' learning opportunities and ensuring that AI-mediated practices do not reproduce or intensify existing social inequalities among students.

Concluding the discussion of the findings in relation to the thesis' third research question, the focus of the analysis has been on the ethical issues that arise in the context of AI-mediated higher education practices. More specifically, to address this question I examined such ethical issues emerging as situated and relational, and arising within specific contexts, stakeholders, practices, and situations rather than as abstract or universal concerns. Aligning with Holmes et al.'s (2022) approach to the ethics of AIED, my analysis has addressed ethical issues in relation to data and education (*data ownership*), algorithms in education (*bias perpetuation*), and the ethics of education itself (*displacement of ethical responsibility in assessment, academic integrity and authorship, fair treatment of students,* and *access to AI*). In addition, drawing on technology mediation, particular attention has been paid to the 'soft impacts' of AGS and GAI chatbots, as perceived and experienced by participants in Studies II–IV. As Kudina (2023) describes, such soft impacts concern technology-inflicted shifts in practices, responsibilities, and value frameworks that are co-constituted and co-evolve with technological mediations.

As with the reconfigurations, the ethical issues identified should not be read as discrete or finite, but as entry points for understanding ethical issues in the context of AI-mediated higher education practices and as part of a shifting ethical landscape that evolves as AI becomes increasingly embedded in teaching and learning, opening new ways of relating with educational practices. Furthermore, the ethical issues outlined in this section should not be perceived as deriving from the technologies themselves but rather are co-constituted by the ways higher education stakeholders relate with AI in the context of existing higher education practices, institutional arrangements, and value systems.

Considered through the lens of technology mediation, the findings from Studies I–IV suggest that the ethical issues – risks, challenges, and concerns – raised in the context of AI-mediated practices in higher education are not entirely new. Instead, they are reconfigured and amplified through the mediating role of the artefacts under study – AGS and GAI chatbots. Long-standing issues related to fairness, academic integrity and equity remain central to higher education, yet the introduction of these technologies reshapes how such issues are understood, negotiated, and enacted in practice, while simultaneously opening new possibilities for action.

## Research contribution

This thesis makes several contributions to the growing body of research on artificial intelligence in higher education (AIHED), artificial intelligence in education (AIED) and AI ethics.

First, it responds to several calls for empirical research in AIHED and AIED, providing empirical insights into an area that has thus far been dominated conceptual debates, policy discussions, and speculative claims about the promises and risks of AI and rather explores how AI technologies are encountered and experienced by university teachers in higher education practices.

Second, the thesis brings ethics to the forefront of AIHED and AIED research, addressing the relative scarcity of studies that critically engage with the ethical dimensions of AI. Additionally, it problematises how ethics is approached in AI and AIED research, arguing for a relational ethics approach to unpack the ethical entanglements of human-technology mediated experiences of practice in higher education. Rather than treating ethics as a fixed list of principles, this thesis unfolds a relational approach, showing how ethics is formed and reformed through AI-mediated practices of negotiation, interpretation, and interaction with the AI artefacts in focus. Furthermore, unlike existing AI ethics literature examining AI mostly in abstract or general terms, this thesis focuses specifically on two technologies – automated grading systems (AGS) and generative AI (GAI) chatbots. This situated focus provides concrete insights into the ways particular technologies mediate educational practices, rather than assuming AI's effects in the abstract.

Third, the thesis brings forward the voices of teachers and illustrates the complexity of their practices in light of emerging AI technologies. While much research in AIHED has focused on students, adoption patterns, or learning outcomes, this thesis foregrounds teachers' experiences. In doing so, it sheds light on how teachers navigate AI-mediated practices, thereby filling a significant gap in the literature.

Finally, the thesis illustrates the complexity of AI mediations in higher education. It challenges instrumentalist understandings of AI as merely a tool

and moves beyond the polarised discourses of technological utopias and dystopias. Instead, it shows how ethics is deeply entangled in these mediations, illustrating the ways in which values, roles and responsibilities are continuously reconfigured as AI technologies become part of educational practice.

# Limitations and delimitations

## Theoretical considerations

This thesis is a postephenomenological study that centres AI mediations in the context of higher education, focusing on AGS and GAI chatbots as examples of AI artefacts and university teachers as key higher education stakeholders. Choosing such an approach means also drawing boundaries in relation to theory. Postphenomenology allows for an account that foregrounds teachers' first-person experiences while also showing how these experiences are technologically mediated, thereby keeping ethics in focus (Rosenberger & Verbeek, 2015). However, a limitation of this choice is that it narrows the analytic gaze of this thesis, limiting its focus on teachers' perspectives and not addressing broader socio-technical structures, institutional discourses, and policy frameworks.

In terms of ethics, the thesis draws on a relational ethics perspective rather than an ethics of care (Noddings, 1984). Relational ethics conceives of ethical meaning as emergent in relations – between humans, technologies, and practices – rather than as a set of abstract principles or as grounded primarily in care and responsibility for the other. This aligns with the postphenomenological commitment to understanding ethics as co-constituted in technological mediations (e.g. fairness, autonomy, and authorship are not stable categories but are constantly reinterpreted in practice as teachers and students engage with AI artefacts). Ethics of care, while highly influential in educational research, emphasises interpersonal attentiveness, empathy, and responsibility (Held, 2005). Although relevant to student-teacher dynamics, it is less suited to theorising how – relatively non-human – technologies mediate ethical relations. The relational ethics perspective better captures the entanglement of human and technological agencies in shaping norms and values. Still, this choice entails limitations as some dimensions of affect, vulnerability, and responsibility emphasised in care ethics may not be fully addressed in this thesis.

### Methodological considerations

Methodologically, this thesis is based on a qualitative research design that foregrounds teachers' voices and lived experiences of AI in higher education. This choice responds to a relative lack of teacher-centred studies in the field, where research has often prioritised students' perspectives, adoption of AI tools, or learning outcomes in connection with intervention studies (Zawacki-Richter et al., 2019; Holmes et al., 2022; McGrath et al., 2025). A limitation of this focus, however, is that it excludes the perspectives of other stakeholders, such as students, administrators, and education policymakers, whose experiences could have enriched the analysis and provided a more multifaceted account of AI mediation in higher education.

The decision to employ qualitative rather than quantitative methods reflects the interest in meaning-making processes, interpretations, and ethical negotiations, rather than in measuring adoption rates or learning outcomes. While this has enabled a rich, practice-near account of teachers' engagements with AI, it also limits the transferability of the findings. Quantitative or mixed-method approaches could complement these insights by examining broader patterns across larger populations or by statistically testing correlations between variables such as attitudes, adoption, and institutional contexts.

Finally, the empirical focus on two specific technologies, AGS and GAI chatbots, offers the strength of concreteness and depth. Studying actual technologies of reference or in use avoids the abstraction that characterises a large body of AIED research. However, this scope also constrains the findings from other forms of AI in higher education (e.g., learning analytics dashboards or adaptive learning platforms). Thus, while AGS and GAI chatbots serve as productive entry points into the broader question of AI mediation in higher education, they cannot fully represent the diversity of technologies currently shaping the field.

## Implications

The thesis suggests that the ethical significance of AI in higher education lies less in the properties of the technologies themselves and more in how they mediate practices, teachers' responsibilities and roles, and student-teacher relationships. Macfarlane's (2003) portrayal of teaching as an ethical practice grounded in professional judgment offers a useful starting point for reflecting on the implications of this thesis in relation to teaching in higher education in light of AI. Like jazz improvisation, teaching unfolds within a shared framework of values, yet requires responsiveness, sensitivity, and attunement to the particularities of each situation. Teaching is a dynamic and context-sensitive

practice in which educators continually balance autonomy with accountability, and rules with compassion. Such a view of what teaching is for and about aligns with Biesta's (2023) conception of teaching as a "double artistry", drawing on Eisner's (2002) and Stenhouse's (1987) descriptions of what education is for and about. Hence, teaching is not a technical activity aimed at producing predetermined outcomes, but a continuous practical art. Education, in this sense, is inherently and unavoidably uncertain (Biesta, 2023; Bonnet & Glazier, 2023b), and if teaching is understood as a "double artistry" of craft and wisdom, then it becomes essential for teachers to continually develop their educational artistry: their capacity to make situated judgments about desirable ways of acting in the always new situations they encounter. Seen from this perspective, what is ethically at stake with AI relates to how it reshapes the conditions under which this double artistry is practised. This thesis illustrated that the emergence of AI unsettles existing practices and reinforces uncertainties and vulnerabilities that characterise educational situations, placing even greater demands on the teachers' judgments. This perspective carries several implications for *future practice*.

First, at the level of policy, it suggests that governance of AI in higher education should move beyond purely technical or compliance-based approaches. Policies should not only regulate AI artefacts but also protect and support the conditions under which teachers can exercise professional judgment. This includes safeguarding spaces for pedagogical experimentation, acknowledging uncertainty as an inherent aspect of educational practice, and avoiding policies that reduce teaching solely to measurable outputs or automated processes. Additionally, current policy discourse on AI in education often lies between utopian promises of innovation and dystopian fears of automation. This thesis suggests the need to move beyond these narratives and engage in more nuanced, practice-informed policymaking. Rather than framing AI adoption solely as a matter of innovation or risk management, policy should attend to the complex, relational, and ethical ways in which AI technologies mediate educational practice and provide resources for institutions to critically engage with these mediations.

Second, at the institutional level, higher education institutions may need to shift from a focus on technological adoption or restriction toward a focus on pedagogical conditions and development. Rather than focusing on which AI systems to procure, deploy, or regulate, institutional strategies could prioritise the cultivation of reflective spaces – forums, seminars, and communities of practice – where educators can collectively interpret the meaning and influences of AI within their disciplines. Professional development, in this light, should not be limited to technical training in how to use AI tools, but should also create opportunities for teachers to reflect on how AI reshapes their roles, responsibilities, and relationships with students. Such initiatives might foreground ethical reflection, pedagogical dialogue, and the cultivation of practical wisdom, rather than presenting AI as a neutral solution to predefined prob-

lems. Additionally, targeted experiential learning activities that engage teachers with practical uses of AI and with the interpretation of institutional regulations may form an important part of this work.

Third, at the departmental level, departments can function as sites for interpreting and reflecting on how AI interacts with disciplinary norms, assessment traditions, and epistemic values. They can help articulate shared principles for appropriate and meaningful uses of AI, while leaving room for individual pedagogical judgment. For teachers as individuals, the presence of AI calls for renewed attention to the judgment-based aspects of their work. Teachers may need to cultivate attentiveness to how AI mediates their practices and reflect on what it highlights, what it obscures, and how it influences their relations with students and conceptualisations of knowledge. This involves ongoing reflection on when to rely on AI, when to resist it, and how to integrate it in ways that remain educationally meaningful. Additionally, at the collective level, teaching communities may play a crucial role in sustaining the 'double artistry' of teaching in higher education. Collegial dialogue, shared experimentation, and collective sense-making can help prevent AI from being reduced to an individualised or purely technical concern, and rather foster shared orientations and support one another in navigating the uncertainties that AI introduces.

Finally, the findings underline the importance of involving teachers and students in the design and deployment of AI systems that are to be used in educational settings. Computer science developers and edtech companies often frame AI solutions in terms of efficiency or scalability, but the mediations identified here demonstrate that pedagogical and ethical dimensions cannot be simply reduced to technical design. Co-design approaches that bring educators and learners into the development process can contribute to designing AI systems that better reflect the needs, values, and lived realities of higher education.

With regard to the implications for *future research* on AI in higher education, first, this thesis illustrates the value of postphenomenology as an approach for examining technologies. Extending this approach to other technologies beyond AGS and GAI chatbots would enrich our understanding of the diverse mediations through which AI reconfigures educational practice.

Second, the thesis highlights the importance of researching ethics as contextually and relationally situated. Ethical challenges are not uniform across settings; they are interpreted, negotiated, and practised differently depending on institutional cultures, disciplinary traditions, and technological infrastructures. Future research should therefore move beyond universal claims about "AI ethics in education" and instead attend to the diversity of contexts in which AI is embedded, foregrounding the specificity of local practices.

Additionally, research on professional development and collective sense-making around AI could include studies exploring how communities of teach-

ers negotiate shared norms, valuing processes, ethical orientations, and practical strategies for integrating AI into their work, and how such collective processes shape the meaning and use of AI in higher education.

Third, the findings call for more longitudinal and practice-based studies that follow how AI technologies are introduced and become embedded in higher education practices as well as and how core practices are re-configured over time. Such research could further explore how roles, responsibilities, and values are gradually reconfigured, and attend to how teachers' and students' needs are changing within AI-mediated practices. Furthermore, there is a need for more empirical studies that explore how AI mediates teaching practices in specific disciplinary and institutional contexts. Such work could examine how teachers interpret and meaningfully integrate AI in their practices and how these mediations shape the space for professional judgment.

Fourth, the thesis suggests future research should explore the potential of AI as a reflective and pedagogical resource. Future studies could investigate how AI technologies might be used to support critical reflection on teaching and assessment practices, for example, as resources for teachers to interrogate their assumptions and understandings of existing practices and engage with ethical dilemmas in educational decision-making and the role of educational judgement.

Finally, further conceptual work is also needed to deepen the understanding of the new pathways of being in and doing education in relation to emerging technologies. This includes examining how concepts such as responsibility, agency, subjectivity, as well as decision-making and praxis – judgment, and action – are being reconfigured through AI mediations, and what this means for the purposes and values of higher education in the years to come.

# Concluding remarks

AI mediations in teaching and learning open new paths for how teachers and students relate to higher education practices. These paths manifest in different ways: through new ways of *doing or not doing* tasks (e.g., AGS automating assessment), new ways of *perceiving* others (e.g., students who use GAI chatbots may be framed as cheaters), and new ways of *relating* to others (e.g., the teacher-AGS-student hermeneutic relation in the interpretation of student work by a teacher using an AGS). Such mediations reconfigure priorities and valuations within educational practice. A telling example is the time-saving promise of AGS, which enables teachers to redirect effort away from assessment toward other pedagogical responsibilities. Here, AGS does not merely reduce workload; it positions higher education stakeholders to actively decide how assessment is valued or devalued – and accordingly its automation is prioritised or downprioritised – in relation to competing demands and responsibilities.

The use and presence of AI in higher education thus give rise to technomoral questions that cannot be reduced to individual choice or intention alone. Questions of authorship, responsibility, fairness, trust, and accountability emerge not simply because of how higher education stakeholders choose to act, but because of how AI technologies mediate educational practices themselves. Much like Verbeek's (2005) example of ultrasound scans reconfiguring parental responsibility through their availability, AI artefacts, such as AGS and GAI chatbots, reconfigure teachers' tasks and responsibilities by making certain forms of teaching, assessment, feedback, and decision-making possible, expected, or unavoidable.

Through the lens of technological mediation, the ethical issues raised by AGS and GAI chatbots are less about introducing wholly new ones than about exposing and amplifying existing ones. Issues such as fairness, academic integrity and equity remain central, but their meanings and enactments shift as these technologies become embedded in educational practice. Following Kudina's (2023) notion of technomoral change, the findings highlight how AI's 'soft impacts' reshape values and responsibilities. Fairness is mediated by tools that claim neutrality yet risk bias, while academic integrity and equity are renegotiated through data-driven infrastructures. In this way, AI reconstitutes the normative frameworks of higher education, transforming how ethics is understood and practised.

This thesis has argued that the ethical significance of AI in higher education lies less in the technologies themselves, and more in the ways they mediate practices, responsibilities, and values. By attending to these mediations, the thesis has positioned AI not as neutral, external, or self-determining, but as something that takes shape within practices and in relation to how stakeholders perceive and interpret the potential beings and doings of AI artefacts. It is through these situated encounters that AI reconfigures how teaching, learning, and assessment are understood and enacted, and how the existential space for praxis – agency, judgment, and action – is opened, constrained, or redirected.

Building on a metaphor of viewing AI like a table lamp introduced into a familiar living room, its significance is shaped by what was already there: the arrangement of the space, the presence of other sources of light, the needs of those inhabiting it, and their willingness to attend to what becomes visible. The room does not look the same after the lamp is switched on, and what emerges is a new state of being in that illuminated space. Yet the lamp does not create objects that were not there before; rather, it makes certain aspects of the room more or less visible under its light. In a similar way, the emergence of AI can illuminate aspects of teaching, learning, and assessment that were previously taken for granted or remained unseen, bringing 'new' tensions, responsibilities, and possibilities into view. Ethical questions thus arise not only from what AI does, but from how it reshapes the conditions under which teachers act, decide, and relate to students.

Seen in this light, the challenge, then, is to cultivate practices and forms of judgment that can sustain meaningful educational action within these evolving sociotechnical conditions (Bearman & Ajjawi, 2023). This thesis offers a step toward such an understanding by foregrounding the mediations through which AI becomes part of higher education, and by opening a space for ongoing reflection on what it means to teach, learn, and act ethically in the light of AI.

# Svensk sammanfattning

Under de senaste åren har framväxten av AI blivit en avgörande fråga för högre utbildning. I såväl interna styrdokument som offentlig debatt framställs AI-utvecklingen i allt högre grad som en nationell angelägenhet, med krav på eller uppmaningar till handling. Samtidigt omformar framväxande AI-teknologier prioriteringar och värderingar inom utbildningspraktiken genom att omforma förutsättningarna för undervisning och lärande. Detta tar sig exempelvis uttryck i nya sätt att utföra eller avstå från uppgifter, och i hur man förstår de olika rollerna och dess respektive ansvar. I detta sammanhang sätter AI-medierade praktiker etiska frågor på agendan – frågor som ofta framställs som helt nya, men som egentligen är djupt rotade i långvariga praktiker inom högre utbildning.

Denna avhandling presenterar en empirisk undersökning av AI-medierade praktiker inom högre utbildning, med särskilt fokus på lärares perspektiv och de etiska frågor som uppstår genom sådana medieringsprocesser. Mer specifikt undersöker jag hur lärare uppfattar och förstår AI-artefakter i relation till sina praktiker, och hur mötet med dessa artefakter omformar deras förståelse av vad de bör göra och hur de bör handla. Detta kan exempelvis vara när lärare ställs inför frågan om hur potentiellt motstridiga krav på autonomi och ansvarsskyldighet ska balanseras. Med utgångspunkt i postfenomenologi och teori om teknologisk mediering fokuserar jag på automatiserade betygssystem (Automated Grading Systems - AGS) och generativa AI (Generative AI - GAI) -chattbotar som framväxande teknologier vars roller, funktioner och betydelser inom- och för högre utbildning ännu inte är etablerade. Teknologiernas rörlighet och snabba utveckling gör dem till tacksamma studieobjekt för att undersöka hur teknologier och högre utbildningspraktiker formar och omformar varandra. Avhandlingen består av fyra sinsemellan kompletterande studier som kombinerar konceptuella och empiriska metoder.

Studie I undersöker etiska aspekter kring AGS med hjälp av relationell etik. Studien går igenom studier om AGS design, utveckling och pilot-studier om AGS i högre utbildning, och analyserar AGS-systemens egenskaper i relation till etiska utmaningar som kan uppstå när de introduceras i högre utbildning. I studien identifieras tre dimensioner av etiska frågor kring AGS: data, algoritmer och bedömning. Studien visar att AGS inte enbart introducerar tekniska

och procedurmässiga överväganden, utan också omformar utbildningspraktiker och relationer på sätt som kräver kontinuerlig, situerad och relationellt medveten etisk reflektion från beslutsfattare.

Studie II är en intervjustudie med AGS-utvecklare som också är universitetslärare och använder dessa system i sin verksamhet. Den undersöker deras förväntningar, erfarenheter, och de störningar som AGS introducerar i deras arbete med bedömning. Resultaten understryker AGS-systemens ambivalenta roll som potentiellt både behjälpliga och lovande, och försvårande och störande. Även om de kan erbjuda effektivitet och förutsägbarhet, introducerar de också nya friktioner och etiska dilemman. AGS framträder inte som neutrala verktyg, utan som sociotekniska system som omformar bedömningsdesign, lärarroller och studentengagemang.

Studie III är en studie inspirerad av det så kallade Turing-testet, följt av fokusgruppsintervjuer med universitetslärare. Studien utforskar hur GAI-chattbotar omformar lärares uppfattningar om deras bedömningspraktiker. Resultaten indikerar att introduktionen av GAI-chattbotar som möjliggör AI-genererat skrivande påverkar lärares bedömningspraktiker och får dem att ifrågasätta författarskap, och att detta i vissa fall också förstärker misstron mellan lärare och studenter, med konsekvenser för relationen dem emellan.
Studie IV är en fokusgruppsstudie som undersöker hur lärare upplever och tolkar framväxten av GAI-chattbotar och hur dessa teknologier omformar lärares förståelse för den egna yrkesrollen. Deltagarna beskrev GAI som både störande och potentiellt transformativ. De föranleddes att ompröva bedömningsformer, undervisningsprioriteringar och sitt eget ansvar i att främja kritisk och etisk förhållningsförmåga till teknologi. Samtidigt föranledde upplevelsen av begränsad beredskap och otillräckligt institutionellt stöd att deltagarna kände sig sårbara i den nya situationen.

De sammantagna resultaten från de fyra studierna visar på att AI:s framväxt rubbar etablerade praktiker och intensifierar de osäkerheter som kännetecknar utbildningssituationer, vilket ställer högre krav på lärares professionella omdömesförmåga. Etiska frågor om författarskap, ansvar, rättvisa, tillit och ansvarsskyldighet uppstår inte enbart ur individuella val, utan ur hur AI-teknologier omformar hur högre utbildning faktiskt görs. Genom att synliggöra dessa omformningar i högre utbildning bidrar avhandlingen till en fördjupande och nyanserande reflektion över vad det innebär att undervisa, lära och handla etiskt inom högre utbildning i ljuset av AI.

# Περίληψη στα ελληνικά

Τα τελευταία χρόνια, η ανάδυση της Τεχνητής Νοημοσύνης (ΤΝ) έχει καταστεί καθοριστικό ζήτημα για την ανώτατη εκπαίδευση παγκοσμίως, και η Σουηδία δεν αποτελεί εξαίρεση. Οι πολιτικές και ακαδημαϊκές αφηγήσεις πλαισιώνουν ολοένα και περισσότερο την Τεχνητή Νοημοσύνης όχι απλώς ως εργαλείο, αλλά ως εθνική επιταγή. Ταυτόχρονα, οι αναδυόμενες τεχνολογίες Τεχνητής Νοημοσύνης αναδιαμορφώνουν τις προτεραιότητες και τις αξιολογήσεις εντός της εκπαιδευτικής πρακτικής, διαμεσολαβώντας τη διδασκαλία και τη μάθηση και ανοίγοντας νέες οδούς για τον τρόπο με τον οποίο εκπαιδευτικοί και φοιτητές σχετίζονται με τις πρακτικές της ανώτατης εκπαίδευσης. Αυτές οι οδοί εκδηλώνονται μέσω νέων τρόπων εκτέλεσης ή μη εκτέλεσης εργασιών, νέων τρόπων αντίληψης των άλλων και νέων τρόπων συσχετισμού με τους άλλους. Σε αυτό το πλαίσιο, οι πρακτικές που διαμεσολαβούνται από την Τεχνητή Νοημοσύνη εγείρουν ηθικά ερωτήματα που συχνά παρουσιάζονται ως πρωτοφανή, αλλά είναι βαθιά ριζωμένα σε μακροχρόνιες πρακτικές της ανώτατης εκπαίδευσης.

Αυτή η διατριβή επιχειρεί μια εμπειρική διερεύνηση των πρακτικών που διαμεσολαβούνται από την Τεχνητή Νοημοσύνη στην ανώτατη εκπαίδευση, φέρνοντας στο προσκήνιο τις προοπτικές των εκπαιδευτικών και εστιάζοντας στα ηθικά ζητήματα που προκύπτουν από τέτοιες μεσολαβήσεις. Πιο συγκεκριμένα, εξετάζω πώς οι πανεπιστημιακοί διδάσκοντες αντιλαμβάνονται και βιώνουν τα τεχνουργήματα της Τεχνητής Νοημοσύνης σε σχέση με τις πρακτικές τους, και πώς αυτά τα τεχνουργήματα μεσολαβούν στην κατανόησή τους για το τι πρέπει να κάνουν και πώς πρέπει να ενεργούν όταν εξισορροπούν συχνά ανταγωνιστικές απαιτήσεις αυτονομίας και λογοδοσίας. Βασιζόμενη στη μεταφαινομενολογία και τη θεωρία της τεχνολογικής διαμεσολάβησης, η διατριβή εστιάζει στα αυτοματοποιημένα συστήματα βαθμολόγησης (automated grading systems – AGS) και στα chatbots γενετικής τεχνητής νοημοσύνης (generative AI chatbots – GAI chatbots) ως αναδυόμενες τεχνολογίες των οποίων οι ρόλοι, οι λειτουργίες και οι έννοιες δεν έχουν ακόμη καθοριστεί. Η ρευστότητά τους τα καθιστά βασικούς χώρους για την εξέταση του τρόπου με τον οποίο οι τεχνολογίες και οι πρακτικές της ανώτατης εκπαίδευσης συνδιαμορφώνουν η μία την άλλη. Η διατριβή είναι μια συλλογή τεσσάρων συμπληρωματικών μελετών που συνδυάζουν εννοιολογικές και εμπειρικές μεθοδολογίες.

Η μελέτη Ι εξετάζει τις ηθικές παραμέτρους των AGS, ανασκοπώντας τη βιβλιογραφία για τα AGS και αναλύοντας τις ιδιαιτερότητές τους μέσω μιας σχεσιακής ηθικής προσέγγισης. Σε αυτήν τη μελέτη, εντοπίζω τρεις διαστάσεις σχετικά με τα ηθικά ζητήματα που περιβάλλουν τα AGS: δεδομένα, αλγόριθμοι και αξιολόγηση. Αυτή η μελέτη τόνισε ότι τα AGS όχι μόνο εισάγουν τεχνικές και διαδικαστικές παραμέτρους, αλλά και αναδιαμορφώνουν τις εκπαιδευτικές πρακτικές και σχέσεις με τρόπους που απαιτούν συνεχή, εντοπισμένη και σχεσιακά ευαίσθητη ηθική αναστοχαστικότητα.

Η μελέτη ΙΙ είναι μια μελέτη βασιζόμενη σε συνέντευξεις με προγραμματιστές AGS, οι οποίοι είναι επίσης πανεπιστημιακοί διδάσκοντες που χρησιμοποιούν αυτά τα συστήματα. Εξετάζει τις προσδοκίες, τις εμπειρίες τους και τις αναταράξεις που εισάγουν τα AGS στις πρακτικές αξιολόγησης. Τα ευρήματα υπογραμμίζουν τον αμφίσημο ρόλο των AGS ως πολλά υποσχόμενων και διαταρακτικών. Ενώ ενδέχεται να προσφέρουν αποτελεσματικότητα και συνέπεια στον τρόπο αξιολόγησης και βαθμολόγησης των εργασιών των φοιτητών, ταυτόχρονα εισάγουν επίσης νέες τριβές και ηθικά διλήμματα. Τα AGS αναδύονται όχι ως ουδέτερα εργαλεία, αλλά ως κοινωνικοτεχνικά συστήματα που αναδιαμορφώνουν τον σχεδιασμό και την οργάνωση της αξιολόγησης, τους ρόλους των διδασκόντων και τη συμμετοχή των μαθητών.

Η μελέτη ΙΙΙ είναι μια πειραματική μελέτη εμπνευσμένη από το τεστ Turing, ακολουθούμενη από συνεντεύξεις ομάδων εστίασης με πανεπιστημιακούς διδάσκοντες. Διερευνά πώς τα GAI chatbots διαμεσολαβούν στις αντιλήψεις των εκπαιδευτικών για τις πρακτικές αξιολόγησής τους. Τα ευρήματα δείχνουν ότι η παρουσία των GAI chatbots επιτρέποντας τη δυνατότητα γραφής που παράγεται από την Τεχνητή Νοημοσύνη διαμορφώνει τις πρακτικές αξιολόγησης, ωθώντας τους εκπαιδευτικούς να αμφισβητήσουν τη συγγραφική ταυτότητα και, σε ορισμένες περιπτώσεις, να ενισχύουν τη δυσπιστία στη σχέση διδάσκοντα–φοιτητή.

Η μελέτη IV είναι μια μελέτη συνεντεύξεων ομάδας εστίασης που εξετάζει πώς οι πανεπιστημιακοί διδάσκοντες βιώνουν και ερμηνεύουν την εμφάνιση των GAI chatbots και πώς αυτές οι τεχνολογίες διαμεσολαβούν στις αντιλήψεις τους για τους επαγγελματικούς τους ρόλους. Οι συμμετέχοντες περιέγραψαν τη γενετική τεχνητή νοημοσύνη ως διαταρακτική και δυνητικά μετασχηματιστική για τις πρακτικές τους. Κλήθηκαν να επανεξετάσουν τις μορφές αξιολόγησης, τις διδακτικές προτεραιότητες και την ευθύνη τους να καλλιεργούν κριτική και ηθική εμπλοκή με την τεχνολογία. Ταυτόχρονα, η περιορισμένη ετοιμότητα και η ανεπαρκής θεσμική υποστήριξη άφησαν τους συμμετέχοντες με αίσθημα ευαλωτότητας

Συνδυαστικά, τα ευρήματα των τεσσάρων μελετών δείχνουν ότι η εμφάνιση της Τεχνητής Νοημοσύνης ανατρέπει τις καθιερωμένες πρακτικές και εντείνει τις αβεβαιότητες που χαρακτηρίζουν τις εκπαιδευτικές

καταστάσεις, θέτοντας μεγαλύτερες απαιτήσεις στην επαγγελματική κρίση των πανεπιστημιακών διδασκόντων. Ηθικά ζητήματα συγγραφικής ταυτότητας, ευθύνης, δικαιοσύνης, εμπιστοσύνης και λογοδοσίας δεν προκύπτουν απλώς από ατομικές επιλογές, αλλά από τον τρόπο με τον οποίο οι τεχνολογίες Τεχνητής Νοημοσύνης διαμεσολαβούν τις πρακτικές της ανώτατης εκπαίδευσης. Αναδεικνύοντας αυτές τις διαμεσολαβήσεις, η διατριβή συμβάλλει στον συνεχιζόμενο αναστοχασμό σχετικά με το τι σημαίνει να διδάσκει, να μαθαίνει και να ενεργεί κανείς ηθικά υπό το φως της Τεχνητής Νοημοσύνης.

# References


Aagaard, J. (2015). Drawn to distraction: A qualitative study of off-task use of educational technology. *Computers & Education*, *87*, 90–97. https://doi.org/10.1016/j.compedu.2015.03.010

Aagaard, J. (2017). Introducing postphenomenological research: A brief and selective sketch of phenomenological research methods. *International Journal of Qualitative Studies in Education*, *30*(6), 519–533. https://doi.org/10.1080/09518398.2016.1263884

Adams, C. A., & Thompson, T. L. (2011). Interviewing objects: Including educational technologies as qualitative research participants. *International Journal of Qualitative Studies in Education*, *24*(6), 733–750. https://doi.org/10.1080/09518398.2010.529849

AI Commission. (2024). *The AI Commission's Roadmap for Sweden* (SOU 2025:12). Statens offentliga utredningar. https://www.sou.gov.se/globalassets/the-ai-commissions-roadmap-for-sweden.pdf

Alexander, K., Savvidou, C., & Alexander, C. (2023). Who Wrote This Essay? Detecting AI-Generated Writing in Second Language Education in Higher Education. *Teaching English with Technology*, *23*(2), 25–43.

Alkhalil, A., Abdallah, M. A. E., Alogali, A., & Aljaloud, A. (2021). Applying Big Data Analytics in Higher Education: A Systematic Mapping Study. *International Journal of Information and Communication Technology Education (IJICTE)*, *17*(3), 29–51. https://doi.org/10.4018/IJICTE.20210701.oa3

ALLEA. (2023). *The European Code of Conduct for Research Integrity*. https://allea.org/code-of-conduct/

Bahja, M., Hammad, R., & Hassouna, M. (2019). Talk2Learn: A Framework for Chatbot Learning. In M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, & J. Schneider (Eds.), *Transforming Learning with Meaningful Technologies* (pp. 582–586). Springer International Publishing. https://doi.org/10.1007/978-3-030-29736-7_44

Bainbridge, L. (1983). Ironies of Automatio. In *Analysis, Design and Evaluation of Man–Machine Systems* (pp. 129–135). Pergamon. https://doi.org/10.1016/B978-0-08-029348-6.50026-9

Baker, T., Smith, L., & Anissa, N. (2019). *Educ-AI-tion Rebooted ? Exploring the future of artificial intelligence in schools and colleges*. Retrieved from nesta. org. uk website: https://media. nesta. org. uk/documents/Future_of_AI_and_education_v5_WEB. pdf.

Bakewell, J. D., Clement-Jones, T. F., Giddens, A., Grender, R. M., Hollick, C. R., Holmes, C., & Levene, P. K. (2018). *AI in the UK: ready, willing and able? .Select committee on artificial intelligence* (pp. 1–183). https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/10002.htm

Balfour, S. P. (2013). Assessing Writing in MOOCs: Automated Essay Scoring and Calibrated Peer Review™. *Research & Practice in Assessment*, *8*, 40–48.

Barker, T. (2011). An Automated Individual Feedback and Marking System: An Empirical Study. *Electronic Journal of E-Learning*, *9*(1), 1–14.

Barnett, R. (2012). Learning for an unknown future. *Higher Education Research & Development*, *31*(1), 65–77. https://doi.org/10.1080/07294360.2012.642841

Barron, A. B., Hebets, E. A., Cleland, T. A., Fitzpatrick, C. L., Hauber, M. E., & Stevens, J. R. (2015). Embracing multiple definitions of learning. *Trends in Neurosciences*, *38*(7), 405–407. https://doi.org/10.1016/j.tins.2015.04.008

Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, *54*(5), 1160–1173. https://doi.org/10.1111/bjet.13337

Bearman, M., & Ajjawi, R. (2024). When I say … artificial intelligence. *Medical Education*, *58*(11), 1273–1275. https://doi.org/10.1111/medu.15408

Bearman, M., Dawson, P., & Tai, J. (2020). Digitally Mediated Assessment in Higher Education: Ethical and Social Impacts. In M. Bearman, P. Dawson, R. Ajjawi, J. Tai, & D. Boud (Eds.), *Re-imagining University Assessment in a Digital World* (pp. 23–36). Springer International Publishing. https://doi.org/10.1007/978-3-030-41956-1_3

Bearman, M., Ryan, J., & Ajjawi, R. (2022). Discourses of artificial intelligence in higher education: A critical literature review. *Higher Education*. https://doi.org/10.1007/s10734-022-00937-2

Bearman, M., Tai, J., Dawson, P., Boud, D., & Ajjawi, R. (2024). Developing evaluative judgement for a time of generative artificial intelligence. *Assessment & Evaluation in Higher Education*, *49*(6), 893–905. https://doi.org/10.1080/02602938.2024.2335321

Bendixen, C., Premat, C., Gunnerstad, A., & Farazouli, A. (2024). *Preventing plagiarism: Handbook for Stockholm University Staff (Second Edition)*. Stockholm University. https://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-235826

Berendt, B. (2019). AI for the Common Good?! Pitfalls, challenges, and ethics pentesting. *Paladyn, Journal of Behavioral Robotics*, *10*(1), 44–65. https://doi.org/10.1515/pjbr-2019-0004

Bergum, V., & Dossetor, J. B. (2005). *Relational Ethics: The Full Meaning of Respect*. University Publishing Group.

Beseiso, M., Alzubi, O. A., & Rashaideh, H. (2021). A novel automated essay scoring approach for reliable higher educational assessments. *Journal of Computing in Higher Education*, *33*(3), 727–746. https://doi.org/10.1007/s12528-021-09283-1

Biesta, G. (2023). Outline of a Theory of Teaching: What Teaching Is, What It Is For, How It Works, and Why It Requires Artistry. In A.-K. Praetorius & C. Y. Charalambous (Eds.), *Theorizing Teaching: Current Status and Open Issues* (pp. 253–280). Springer International Publishing. https://doi.org/10.1007/978-3-031-25613-4_9

Birhane, A. (2021). Algorithmic injustice: A relational ethics approach. *Patterns*, *2*(2), 100205. https://doi.org/10.1016/j.patter.2021.100205

Bommarito, M. J., & Katz, D. M. (2022). *GPT Takes the Bar Exam* (SSRN Scholarly Paper No. 4314839). https://doi.org/10.2139/ssrn.4314839

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 in-

telligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, *21*(1), Article 1. https://doi.org/10.1186/s41239-023-00436-z

Bonnet, A., & Glazier, J. (2023b). The conflicted role of uncertainty in teaching and teacher education. *Teachers and Teaching*, *31*(2), 201–217. https://doi.org/10.1080/13540602.2023.2272650

Bonnet, A., & Glazier, J. (2024a). Educating for uncertainty in what is certainly an uncertain world. *Teachers and Teaching*, *31*(2), 168–173. https://doi.org/10.1080/13540602.2024.2320160

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2021). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative Research in Sport, Exercise and Health*, *13*(2), 201–216. https://doi.org/10.1080/2159676X.2019.1704846

Brooks, M., Basu, S., Jacobs, C., & Vanderwende, L. (2014). Divide and correct: Using clusters to grade short answers at scale. *Proceedings of the First ACM Conference on Learning @ Scale Conference, L@S '14*, 89–98. https://doi.org/10.1145/2556325.2566243

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). *Language Models are Few-Shot Learners* (arXiv:2005.14165). arXiv. https://doi.org/10.48550/arXiv.2005.14165

Buchanan, C., Howitt, M. L., Wilson, R., Booth, R. G., Risling, T., & Bamford, M. (2021). Predicted Influences of Artificial Intelligence on Nursing Education: Scoping Review. *JMIR Nursing*, *4*(1), e23933. https://doi.org/10.2196/23933

Carlson, L. (2016). *Academic Freedom and Rights to University Teaching Materials: A Comparison of Swedish, American and German Approaches* (SSRN Scholarly Paper No. 2713421). Social Science Research Network. https://doi.org/10.2139/ssrn.2713421

Caukin, N., & Brinthaupt, T. (2017). Using a Teaching Philosophy Statement as a Professional Development Tool for Teacher Candidates. *International Journal for the Scholarship of Teaching and Learning*, *11*(2). https://doi.org/10.20429/ijsotl.2017.110218

Cerratto Pargman, T., & McGrath, C. (2021). Mapping the Ethics of Learning Analytics in Higher Education: A Systematic Literature Review of Empirical Research. *Journal of Learning Analytics*, *8*(2), 123–139. https://doi.org/10.18608/jla.2021.1

Cerratto Pargman, T., McGrath, C., Viberg, O., & Knight, S. (2023). New Vistas on Responsible Learning Analytics: A Data Feminist Perspective. *Journal of Learning Analytics*, *10*(1), Article 1. https://doi.org/10.18608/jla.2023.7781

Cerratto Pargman, T., Sporrong, E., Farazouli, A., & McGrath, C. (2024). Beyond the Hype: Towards a Critical Debate About AI Chatbots in Swedish Higher Education. *Högre Utbildning*, *14*(1), Article 1. https://doi.org/10.23865/hu.v14.6243

Chen, X., Xie, H., Zou, D., & Hwang, G.-J. (2020). Application and theory gaps during the rise of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, *1*, 100002. https://doi.org/10.1016/j.caeai.2020.100002

Chia, R., & Holt, R. (2008). On Managerial Knowledge. *Management Learning*, *39*(2), 141–158. https://doi.org/10.1177/1350507607087579

Choi, J. H., Hickman, K. E., Monahan, A., & Schwarcz, D. (2023). *ChatGPT Goes to Law School* (SSRN Scholarly Paper No. 4335905). Social Science Research Network. https://doi.org/10.2139/ssrn.4335905

Christensen, C. M., Raynor, M. E., & McDonald, R. (2015). What Is Disruptive Innovation? *Harvard Business Review*. https://hbr.org/2015/12/what-is-disruptive-innovation

Coeckelbergh, M. (2019). *Moved by Machines: Performance Metaphors and Philosophy of Technology*. Routledge. https://doi.org/10.4324/9780429283130

Coeckelbergh, M. (2020). Technoperformances: Using metaphors from the performance arts for a postphenomenology and posthermeneutics of technology use. *AI & SOCIETY*, *35*(3), 557–568. https://doi.org/10.1007/s00146-019-00926-7

Colonna, L. (2024). Teachers in the loop? An analysis of automatic assessment systems under Article 22 GDPR. *International Data Privacy Law*, *14*(1), 3–18. https://doi.org/10.1093/idpl/ipad024

Corbin, T., Dawson, P., Nicola-Richmond, K., & Partridge, H. (2025). 'Where's the line? It's an absurd line': towards a framework for acceptable uses of AI in assessment. *Assessment & Evaluation in Higher Education*, *50*(5), 705–717. https://doi.org/10.1080/02602938.2025.2456207

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. In *Innovations in Education and Teaching International* (Vol. 61, Issue 2, pp. 228–239). https://doi.org/10.1080/14703297.2023.2190148

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, *20*(1), 22. https://doi.org/10.1186/s41239-023-00392-8

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*, 100179. https://doi.org/10.1016/j.caeai.2023.100179

Decuypere, M., & Simons, M. (2016). Relational thinking in education: Topology, sociomaterial studies, and figures. *Pedagogy, Culture & Society*, *24*(3), 371–386. https://doi.org/10.1080/14681366.2016.1166150

Dignum, V. (2019). *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Springer International Publishing. https://doi.org/10.1007/978-3-030-30371-6

Dikli, S. (2010). The Nature of Automated Essay Scoring Feedback. *CALICO Journal*, *28*(1), 99–134.

Dwyer, S. C., & Buckle, J. L. (2009). The Space Between: On Being an Insider-Outsider in Qualitative Research. *International Journal of Qualitative Methods*, *8*(1), 54–63. https://doi.org/10.1177/160940690900800105

Eisenberg, M., & Fischer, G. (2014). MOOCs: A perspective from the learning sciences. *Proceedings of International Conference of the Learning Sciences, ICLS*, *1*, 190–197.

Eisner, E. W. (2002). From episteme to phronesis to artistry in the study and improvement of teaching. *Teaching and Teacher Education*, *18*(4), 375–385. https://doi.org/10.1016/S0742-051X(02)00004-5

Ekberg, T. (2025). *Akademisk frihet så in i Norden: En studie om hur akademisk frihet och institutionell autonomi har utvecklats för universitet och högskolor i Danmark, Finland och Norge samt förslag hur detta kan öka i Sverige*. SUHF [Association of Swedish Higher Education Institutions].

Ekberg, T., & Söderbergh Widding, A. (2025). *En kodex för universitet och högskolor: Grundläggande principer och kärnvärden*. SUHF [Association of Swedish Higher Education Institutions].

Elkins, K., & Chun, J. (2020). Can GPT-3 Pass a Writer's Turing Test? *Journal of Cultural Analytics*, *5*. https://doi.org/10.22148/001c.17212

Erhardt, C., Kullenberg, H., Grigoriadis, A., Kumar, A., Christidis, N., & Christidis, M. (2025). From policy to practice: The regulation and implementation of generative AI in Swedish higher education institutes. *International Journal for Educational Integrity*, *21*(1), 21. https://doi.org/10.1007/s40979-025-00195-6

European Commission. (2022, October 25). *Ethical guidelines on the use of artificial intelligence and data in teaching and learning for educators | European Education Area*. https://education.ec.europa.eu/node/2285

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gasevic, D. (2024). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *BRITISH JOURNAL OF EDUCATIONAL TECHNOLOGY*. https://doi.org/10.1111/bjet.13544

Farazouli, A. (2024). Automation and Assessment: Exploring Ethical Issues of Automated Grading Systems from a Relational Ethics Approach. In A. Buch, Y. Lindberg, & T. Cerratto Pargman (Eds.), *Framing Futures in Postdigital Education: Critical Concepts for Data-driven Practices* (pp. 209–226). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-58622-4_12

Farazouli, A., Cerratto Pargman, T., Bolander Laksov, K., & McGrath, C. (2025). Navigating uncertainty: University teachers' experiences and perceptions of generative artificial intelligence in teaching and learning. *Studies in Higher Education*, *0*(0), 1–16. https://doi.org/10.1080/03075079.2025.2550766

Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2024). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment and Evaluation in Higher Education*, *49*(3), 363–375. https://doi.org/10.1080/02602938.2023.2241676

Feigenbaum, E. A. (2003). Some challenges and grand challenges for computational intelligence. *Journal of the ACM*, *50*(1), 32–40. https://doi.org/10.1145/602382.602400

Figueras, C., Farazouli, A., Cerratto Pargman, T., McGrath, C., & Rossitto, C. (2025). Promises and breakages of automated grading systems: A qualitative study in computer science education. *Education Inquiry*, *0*(0), 1–22. https://doi.org/10.1080/20004508.2025.2464996

Fisher, E. (2010). Contemporary Technology Discourse and the Legitimation of Capitalism. *European Journal of Social Theory*, *13*(2), 229–252. https://doi.org/10.1177/1368431010362289

Flodén, J. (2025). Grading exams using large language models: A comparison between human and AI grading of exams in higher education using ChatGPT. *British Educational Research Journal*, *51*(1), 201–224. https://doi.org/10.1002/berj.4069

Floridi, L. (2019). Translating Principles into Practices of Digital Ethics: Five Risks of Being Unethical. *Philosophy & Technology*, *32*(2), 185–193. https://doi.org/10.1007/s13347-019-00354-x

Forsberg, E., Ziegert, K., Hult, H., & Fors, U. (2015). Evaluation of a novel scoring and grading model for VP-based exams in postgraduate nurse education. *Nurse Education Today*, *35*(12), 1246–1251. https://doi.org/10.1016/j.nedt.2015.04.005

Generative Pretrained Transformer, G., Thunström, A. O., & Steingrimsson, S. (2022). *Can GPT-3 write an academic paper on itself, with minimal human input?* https://hal.science/hal-03701250

Gierl, M. J., Latifi, S., Lai, H., Boulais, A.-P., & De Champlain, A. (2014). Automated essay scoring and the future of educational assessment in medical education. *Medical Education*, *48*(10), 950–962. https://doi.org/10.1111/medu.12517

Gilliard, C., & Rorabaugh, P. (2023, February 3). You're Not Going to Like How Colleges Respond to ChatGPT. *Slate*. https://slate.com/technology/2023/02/chatgpt-cheating-college-ai-detection.html

Giray, L., Sevnarayan, K., & Ranjbaran Madiseh, F. (2025). Beyond Policing: AI Writing Detection Tools, Trust, Academic Integrity, and Their Implications for College Writing. *Internet Reference Services Quarterly*, *29*(1), 83–116. https://doi.org/10.1080/10875301.2024.2437174

GMI. (2023). *AI in Education Market Statistics, Trends & Growth Opportunity 2032*. Global Market Insights Inc. https://www.gminsights.com/industry-analysis/artificial-intelligence-ai-in-education-market

Graneheim, U. H., & Lundman, B. (2004). Qualitative content analysis in nursing research: Concepts, procedures and measures to achieve trustworthiness. *Nurse Education Today*, *24*(2), 105–112. https://doi.org/10.1016/j.nedt.2003.10.001

Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: Secrets of the trade. *Journal of Chiropractic Medicine*, *5*(3), 101–117. https://doi.org/10.1016/S0899-3467(07)60142-6

Hagendorff, T. (2020). The Ethics of AI Ethics: An Evaluation of Guidelines. *Minds and Machines*, *30*(1), 99–120. https://doi.org/10.1007/s11023-020-09517-8

Harland, T., & Pickering, N. (2011). *Values in Higher Education Teaching*. Routledge. https://doi.org/10.4324/9780203842003

Heidegger, M. (with Internet Archive). (1977). *The question concerning technology, and other essays*. New York : Harper Torchbooks. http://archive.org/details/questionconcerni00heid

Held, V. (2005). *The Ethics of Care: Personal, Political, and Global*. Oxford University Press. https://doi.org/10.1093/0195180992.001.0001

Henderson, M., Bearman, M., Chung, J., Fawns, T., Buckingham Shum, S., Matthews, K. E., & de Mello Heredia, J. (2025). Comparing Generative AI and teacher feedback: Student perceptions of usefulness and trustworthiness. *Assessment & Evaluation in Higher Education*, *0*(0), 1–16. https://doi.org/10.1080/02602938.2025.2502582

Herjevik, M. (2017). *UKÄ 2017: Rättssäker examination—Tredje upplagan*. Universitetskanslersämbetet. https://www.uka.se/download/18.16cf0f8c1849df46622163/1669103511964/2017-rattssaker-examination.pdf

Herjevik, M. (2020). *UKÄ 2020: Rättssäker examination Fjärde upplagan* (Nos. 32-314–18). Universitetskanslersämbetet. https://www.uka.se/download/18.16cf0f8c1849df46622152/1669103146069/Vagledning-2020-01-16-rattssaker-examination.pdf

Herodotou, C., Maguire, C., McDowell, N., Hlosta, M., & Boroowa, A. (2021). The engagement of university teachers with predictive learning analytics. *Computers & Education*, *173*, 104285. https://doi.org/10.1016/j.compedu.2021.104285

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

Hillebrand, B., Driessen, P. H., & Koll, O. (2015). Stakeholder marketing: Theoretical foundations and required capabilities. *Journal of the Academy of Marketing Science*, *43*(4), 411–428. https://doi.org/10.1007/s11747-015-0424-y

Högskoleförordningen [Higher Education Ordinance], 1993:100; amended by Förordning 2021:1338 (1993). https://www.riksdagen.se/sv/dokument-och-lagar/dokument/svensk-forfattningssamling/hogskoleforordning-1993100_sfs-1993-100/

Holloway, R. (with Internet Archive). (1999). *Godless morality: Keeping religion out of ethics*. Edinburgh : Canongate. http://archive.org/details/godlessmoralityk0000holl

Holmes, W. (2021). *AI and Education: A Critical Studies Perspective. Utilizing AI in Developing Education Systems* [Working / discussion paper]. UNESCO RCEP (Regional Center for Educational Planning). UNESCO RCEP (Regional Center for Educational Planning): Sharjah, UAE. (2021). https://rcepunesco.ae/en/KnowledgeCorner/WorkingPapers/WorkingPapers/2021%20-%20Wayne%20Holmes%20-%20AI%20and%20Education-%20A%20Critical%20Studies%20Perspective.pdf

Holmes, W., Mouta, A., Hillman, V., Schiff, D., Laak, K.-J., Atenas, J., Bardone, E., Lochead, K., Gonsales, P., Havemann, L., Seon, J., Go, B., Schreurs, B., Zhgenti, S., Lee, K., Bali, M., Bialik, M., Medina-Gual, L., Knight, S., … Yeo, B. (2025). *Critical Studies of Artificial Intelligence and Education: Putting a Stake in the Ground.* (SSRN Scholarly Paper No. 5391793). Social Science Research Network. https://doi.org/10.2139/ssrn.5391793

Holmes, W., & Porayska-Pomsta, K. (Eds.). (2022). *The Ethics of Artificial Intelligence in Education: Practices, Challenges, and Debates*. Routledge. https://doi.org/10.4324/9780429329067

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in Education: Towards a Community-Wide Framework. *International Journal of Artificial Intelligence in Education*, *32*(3), 504–526. https://doi.org/10.1007/s40593-021-00239-1

Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, *57*(4), 542–570. https://doi.org/10.1111/ejed.12533

Ihde, D. (1990). *Technology and the lifeworld: From garden to earth*. Indiana University Press.

Ihde, D. (1993). *Postphenomenology: Essays in the postmodern context*. Northwestern University Press.

Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review*, *10*(1), 18–26. https://doi.org/10.1007/s13162-020-00161-0

Jackson, S. J. (2014). Rethinking Repair. In T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), *Media Technologies: Essays on Communication, Materiality, and Society* (p. 0). The MIT Press. https://doi.org/10.7551/mitpress/9780262525374.003.0011

Jackson, S. J. (2016). Speed, Time, Infrastructure: Temporalities of Breakdown, Maintenance, and Repair. In J. Wajcman & N. Dodd (Eds.), *The Sociology of Speed: Digital, Organizational, and Social Temporalities* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198782858.003.0012

Janssen, A., Grützner, L., & Breitner, M. (2021). *Why do Chatbots fail? A Critical Success Factors Analysis*. International Conference on Interaction Sciences. https://www.semanticscholar.org/paper/Why-do-Chatbots-fail-A-Critical-Success-Factors-Janssen-Gr%C3%BCtzner/5cb8bcd29aca1849bdb0e22972d6b1cb6f70979f

Jasanoff, S., & Kim, S.-H. (2009). Containing the Atom: Sociotechnical Imaginaries and Nuclear Power in the United States and South Korea. *Minerva*, *47*(2), 119–146. https://doi.org/10.1007/s11024-009-9124-4

Jasanoff, S., & Kim, S.-H. (Eds.). (2015). *Dreamscapes of Modernity: Sociotechnical Imaginaries and the Fabrication of Power*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226276663.001.0001

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, *1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Jurafsky, D., & Martin, J. H. (2023, January). *Speech and Language Processing*. https://web.stanford.edu/~jurafsky/slp3/

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). *ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education*. EdArXiv. https://doi.org/10.35542/osf.io/5er8f

Kemmis, S. (2022). The Theory of Practice Architectures: Practices. In S. Kemmis (Ed.), *Transforming Practices: Changing the World with the Theory of Practice Architectures* (pp. 53–74). Springer Nature. https://doi.org/10.1007/978-981-16-8973-4_4

Kizilcec, R. F., & Lee, H. (2022). Algorithmic fairness in education. In *The Ethics of Artificial Intelligence in Education*. Routledge.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence*, *5*. https://doi.org/10.1016/j.caeai.2023.100156

Kudina, O. (2023). *Moral Hermeneutics and Technology: Making Moral Sense through Human-Technology-World Relations*. Lexington Books.

Kyrk, P., Ellfolk Kenttä, E., & Axelsson, S. (2024). *UKÄ 2025:7 Disciplinärenden 2024 vid universitet och högskolor* (2025:7). Universitetskanslersämbetet. https://www.uka.se/download/18.92baa461992cb98f7614/1757583812354/Disciplin%C3%A4renden%202024%20vid%20universitet%20och%20hogskolor.pdf

Kyrk, P., Viberg, A., & Axelsson, S. (2023). *UKÄ 2024:4 Disciplinärenden 2023 vid universitet och högskolor* (2024:4). Universitetskanslersämbetet. https://www.uka.se/download/18.2787816f18e2141329f2a745/1710175096770/Disciplin%C3%A4renden%202023%20vid%20universitet%20och%20h%C3%B6gskolor.pdf

Larkey, L. S. (1998). Automatic essay grading using text categorization techniques. *Proceedings of the 21st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '98*, 90–95. https://doi.org/10.1145/290941.290965

Leacock, C., & Chodorow, M. (2003). C-rater: Automated Scoring of Short-Answer Questions. *Computers and the Humanities*, *37*(4), 389–405. https://doi.org/10.1023/A:1025779619903

Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F., Lekkas, D., & 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*, 100221. https://doi.org/10.1016/j.caeai.2024.100221

Long, D., & Magerko, B. (2020). What is AI Literacy? Competencies and Design Considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, CHI '20*, 1–16. https://doi.org/10.1145/3313831.3376727

Luckin, R., Cukurova, M., Kent, C., & du Boulay, B. (2022). Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence*, *3*, 100076. https://doi.org/10.1016/j.caeai.2022.100076

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence Unleashed: An argument for AI in Education. In *UCL Knowledge Lab: London, UK.* [Report]. UCL Knowledge Lab. https://www.pearson.com/content/dam/corporate/global/pearson-dot-com/files/innovation/Intelligence-Unleashed-Publication.pdf

Luo, J. (Jess), & Dawson, P. (2025). Exploring value judgements in grading: Will teachers mark down student work assisted by GenAI, and should they? *Studies in Higher Education*, *0*(0), 1–15. https://doi.org/10.1080/03075079.2025.2552825

Luo (Jess), J. (2025). How does GenAI affect trust in teacher-student relationships? Insights from students' assessment experiences. *Teaching in Higher Education*, *30*(4), 991–1006. https://doi.org/10.1080/13562517.2024.2341005

Macfarlane, B. (2003). *Teaching with Integrity: The Ethics of Higher Education Practice*. Routledge. https://doi.org/10.4324/9780203416501

Maldonado-Mahauad, J., Pérez-Sanagustín, M., Carvallo-Vega, J., Narvaez, E., & Calle, M. (2022). Miranda: A Chatbot for Supporting Self-regulated Learning. In I. Hilliger, P. J. Muñoz-Merino, T. De Laet, A. Ortega-Arranz, & T. Farrell (Eds.), *Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption* (pp. 455–462). Springer International Publishing. https://doi.org/10.1007/978-3-031-16290-9_36

Maslej, N., Fattorini, L., Raymond Perrault, Yolanda Gil, Vanessa Parli, Njenga Kariuki, Emily Capstick, Anka Reuel, Erik Brynjolfsson, John Etchemendy, Katrina Ligett, Terah Lyons, James Manyika, Juan Carlos Niebles, Yoav Shoham, Russell Wald, Tobi Walsh, Armin Hamrah, Lapo Santarlasci, … Sukrut Oak. (2025). *The AI Index 2025 Annual Report*. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. https://hai.stanford.edu/ai-index/2025-ai-index-report

Mayumi Takaya, Yusuke Tsuruta, & Akihiro Yamamura. (2013). Reverse Turing Test using Touchscreens and CAPTCHA. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, *4*(3), 41–57. https://doi.org/10.22667/JOWUA.2013.09.31.041

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology*, *19*(4–5), 494–513. https://doi.org/10.1080/09541440701326154

McGrath, C., & Åkerfeldt, A. (2019). Educational technology (EdTech): Unbounded opportunities or just another brick in the wall? In *Digital Transformation and Public Services*. Routledge.

McGrath, C., Cerratto Pargman, T., Juth, N., & Palmgren, P. J. (2023). University teachers' perceptions of responsibility and artificial intelligence in higher education—An experimental philosophical study. *Computers and Education: Artificial Intelligence*, *4*, 100139. https://doi.org/10.1016/j.caeai.2023.100139

McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2025). Generative AI chatbots in higher education: A review of an emerging research area. *Higher Education*, *89*(6), 1533–1549. https://doi.org/10.1007/s10734-024-01288-w

McGrath, C., Stenfors-Hayes, T., Roxå, T., & Bolander Laksov, K. (2017). Exploring dimensions of change: The case of MOOC conceptions. *International Journal for Academic Development*, *22*(3), 257–269. https://doi.org/10.1080/1360144X.2017.1291430

McNamara, A., Smith, J., & Murphy-Hill, E. (2018). Does ACM's code of ethics change ethical decision making in software development? *Proceedings of the 2018 26th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering, ESEC/FSE 2018*, 729–733. https://doi.org/10.1145/3236024.3264833

Miao, F., Holmes, W., Ronghuai, H., & Hui, Z. (2021). *AI and education: Guidance for policy-makers*. UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000376709

Microsoft. (2023, April 20). *Responsible and trusted AI - Cloud Adoption Framework*. https://learn.microsoft.com/en-us/azure/cloud-adoption-framework/innovate/best-practices/trusted-ai

Mirzababaei, B., & Pammer-Schindler, V. (2022). An Educational Conversational Agent for GDPR. In I. Hilliger, P. J. Muñoz-Merino, T. De Laet, A. Ortega-Arranz, & T. Farrell (Eds.), *Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption* (pp. 470–476). Springer International Publishing. https://doi.org/10.1007/978-3-031-16290-9_38

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, *1*(11), 501–507. https://doi.org/10.1038/s42256-019-0114-4

Morell-Mengual, V., Fernández-García, O., Berenguer, C., Ortega-Barón, J., Gil-Llario, M. D., & Estruch-García, V. (2025). Characteristics, motivations and attitudes of students using ChatGPT and other language model-based chatbots in higher education. *Education and Information Technologies*, *30*(15), 22257–22274. https://doi.org/10.1007/s10639-025-13650-1

Moss, G., & Jewitt, C. (2010). Policy, Pedagogy and Interactive Whiteboards: What Lessons Can be Learnt from Early Adoption in England? In *Interactive Whiteboards for Education: Theory, Research and Practice* (pp. 20–36). IGI Global. (policy-pedagogy-interactive-whiteboards). https://doi.org/10.4018/978-1-61520-715-2.ch002

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences*, *116*(44), 22071–22080. https://doi.org/10.1073/pnas.1900654116

Nguyen, T. C. (2023). University Teachers' Perceptions of Using ChatGPT in Language Teaching and Assessment. *Proceedings of the AsiaCALL International Conference*, *4*, 116–128. https://doi.org/10.54855/paic.2349

Nicolini, D. (2009). Zooming In and Out: Studying Practices by Switching Theoretical Lenses and Trailing Connections. *Organization Studies*, *30*(12), 1391–1418. https://doi.org/10.1177/0170840609349875

Nicolini, D. (2013). *Practice Theory, Work, and Organization: An Introduction*. Oxford University Press.

Noddings, N. (1984). *Caring, a feminine approach to ethics & moral education*. University of California Press.

Odelberg, N., Ärnhage, M., & Nordström, K. (2023). *UKÄ 2023: Artificiell intelligens och högskolans utbildningsutbud* (Redovisning av ett regeringsuppdrag UKÄ 2023:). Universitetskanslersämbetet. https://www.uka.se/download/18.36bb9e318e560741e38d/1712214483255/Artificiell%20intelligens%20och%20h%C3%B6gskolans%20utbildningsutbud.pdf

OECD. (2021). *OECD Digital Education Outlook 2021: Pushing the Frontiers with Artificial Intelligence, Blockchain and Robots*. OECD. https://doi.org/10.1787/589b283f-en

OECD. (2023). *OECD Digital Education Outlook 2023: Towards an Effective Digital Education Ecosystem*. OECD Publishing. https://doi.org/10.1787/c74f03de-en

OpenAI. (2023). *GPT-4 Technical Report* (arXiv:2303.08774). arXiv. https://doi.org/10.48550/arXiv.2303.08774

Pardo, A., & Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, *45*(3), 438–450. https://doi.org/10.1111/bjet.12152

Perkins, M., & Roe, J. (2024). Decoding Academic Integrity Policies: A Corpus Linguistics Investigation of AI and Other Technological Threats. In *HIGHER EDUCATION POLICY* (Vol. 37, Issue 3, pp. 633–653). PALGRAVE MACMILLAN LTD. https://doi.org/10.1057/s41307-023-00323-2

Pillow, W. (2003). Confession, catharsis, or cure? Rethinking the uses of reflexivity as methodological power in qualitative research. *International Journal of Qualitative Studies in Education*, *16*(2), 175–196. https://doi.org/10.1080/0951839032000060635

Pink, S. (2019). Digital Social Futures Research. *Journal of Digital Social Research*, *1*(1), 41–48. https://doi.org/10.33621/jdsr.v1i1.13

Pink, S. (2022). *Emerging Technologies / Life at the Edge of the Future*. Routledge. https://doi.org/10.4324/9781003182528

Pollard, C. (2015). *What is the right thing to do: Use of a relational ethic framework to guide clinical decision-making*. https://www.semanticscholar.org/paper/What-is-the-right-thing-to-do%3A-use-of-a-relational-Pollard/59eb729d4ceff678cb1255b18505954b12cf22c7

Porter, T. M. (2020). Cultures of Objectivity. In T. M. Porter (Ed.), *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life* (p. 0). Princeton University Press. https://doi.org/10.23943/princeton/9780691208411.003.0001

Premat, C., & Farazouli, A. (2025). Academic Integrity vs. Artificial Intelligence: A tale of two AIs. *Práxis Educativa*, *20*, 1–12. https://doi.org/10.5212/PraxEduc.v.20.24871.016

Pressey, S. L. (1926). A simple device for teaching, testing, and research in learning. *School and Society*, *23*, 373–376.

Pulman, S. G., & Sukkarieh, J. Z. (2005). Automatic short answer marking. *Proceedings of the Second Workshop on Building Educational Applications Using NLP, EdAppsNLP 05*, 9–16.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). *Language Models are Unsupervised Multitask Learners*. https://www.semanticscholar.org/paper/Language-Models-are-Unsupervised-Multitask-Learners-Radford-Wu/9405cc0d6169988371b2755e573cc28650d14dfe

Rosenberger, R. (2012). Embodied technology and the dangers of using the phone while driving. *Phenomenology and the Cognitive Sciences*, *11*(1), 79–94. https://doi.org/10.1007/s11097-011-9230-2

Rosenberger, R., & Verbeek, P. P. (2015). *Postphenomenological investigations: Essays on human-technology relations*. Lexington Books.

Ross, J., Sinclair, C., Knox, J., Bayne, S., & Macleod, H. (2014). Teacher Experiences and Academic Identity: The Missing Components of MOOC Pedagogy. *Journal of Online Learning and Teaching*, *10*(1), 57–69.

Rotolo, D., Hicks, D., & Martin, B. R. (2015). What is an emerging technology? *Research Policy*, *44*(10), 1827–1843. https://doi.org/10.1016/j.respol.2015.06.006

Rumrill, P. D., & Fitzgerald, S. M. (2001). Using narrative literature reviews to build a scientific knowledge base. *Work (Reading, Mass.)*, *16*(2), 165–170.

Russell, S. J., Norvig, P., & Davis, E. (2010). *Artificial Intelligence: A Modern Approach*. Prentice Hall.

Saunders, S. (2023, February 24). *Rather than ban generative AI, universities must learn from the past*. University World News. https://www.universityworldnews.com/post.php?story=20230221075136410

Scao, T. L., Fan, A., Akiki, C., Pavlick, E., Ilić, S., Hesslow, D., Castagné, R., Luccioni, A. S., Yvon, F., Gallé, M., Tow, J., Rush, A. M., Biderman, S., Webson, A., Ammanamanchi, P. S., Wang, T., Sagot, B., Muennighoff, N., del Moral, A. V., … Wolf, T. (2023). *BLOOM: A 176B-Parameter Open-Access Multilingual Language Model* (arXiv:2211.05100). arXiv. https://doi.org/10.48550/arXiv.2211.05100

Schiff, D. (2022). Education for AI, not AI for Education: The Role of Education and Ethics in National AI Policy Strategies. *International Journal of Artificial Intelligence in Education*, *32*(3), 527–563. https://doi.org/10.1007/s40593-021-00270-2

Sclater, N. (2016). Developing a Code of Practice for Learning Analytics. *Journal of Learning Analytics*, *3*(1), 16–42. https://doi.org/10.18608/jla.2016.31.3

Selinger, E. (Ed.). (2006). *Postphenomenology: A critical companion to Ihde*. State University of New York Press.

Selwyn, N. (2016). *Is Technology Good for Education?* Polity Press. https://research.monash.edu/en/publications/is-technology-good-for-education

Selwyn, N. (2022). AI, education, and ethics – starting a conversation. In *The Ethics of Artificial Intelligence in Education—Practices, Challenges, and Debates* (1st ed., p. 312). Routledge. https://doi.org/10.4324/9780429329067

Selwyn, N., Hillman, T., Bergviken Rensfeldt, A., & Perrotta, C. (2023). Digital Technologies and the Automation of Education—Key Questions and Concerns. *Postdigital Science and Education*, *5*(1), 15–24. https://doi.org/10.1007/s42438-021-00263-3

Selwyn, N., Ljungqvist, M., & Sonesson, A. (2025). When the prompting stops: Exploring teachers' work around the educational frailties of generative AI tools. *Learning, Media and Technology*, *50*(3), 310–323. https://doi.org/10.1080/17439884.2025.2537959

Sergeeva, A. V. (2023). A Postphenomenological Perspective On the Changing Nature of Work. *Computer Supported Cooperative Work (CSCW)*, *32*(2), 215–236. https://doi.org/10.1007/s10606-022-09447-2

Shay, B. S. (2008). The Assessment of Complex Performance: A Socially Situated Interpretive Act. *Harvard Educational Review*, *74*(3), 307–329. https://doi.org/10.17763/haer.74.3.wq16l67103324520

Siemens, G., & Baker, R. S. J. d. (2012). Learning analytics and educational data mining: Towards communication and collaboration. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, LAK '12*, 252–254. https://doi.org/10.1145/2330601.2330661

Sismondo, S. (2020). Sociotechnical imaginaries: An accidental themed issue. *Social Studies of Science*, *50*(4), 505–507. https://doi.org/10.1177/0306312720944753

Skinner, B. F. (1958). Teaching machines; from the experimental study of learning come devices which arrange optimal conditions for self instruction. *Science (New York, N.Y.)*, *128*(3330), 969–977. https://doi.org/10.1126/science.128.3330.969

Slade, S., & Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist*, *57*(10), 1510–1529. https://doi.org/10.1177/0002764213479366

Slattery, P., & Rapp, D. (2002). *Ethics and the Foundations of Education: Teaching Convictions in a Postmodern World* (1st edition). Allyn & Bacon.

Sporrong, E., McGrath, C., & Cerratto Pargman, T. (2025). Situating AI in assessment—An exploration of university teachers' valuing practices. *AI and Ethics*, *5*(3), 2381–2394. https://doi.org/10.1007/s43681-024-00558-8

Sporrong, E., McGrath, C., Viberg, O., & Cerratto Pargman, T. (2025). What is the problem with generative artificial intelligence in higher education? – A critical analysis of educator responsibility in the Swedish policy landscape. *Learning, Media and Technology*, *0*(0), 1–20. https://doi.org/10.1080/17439884.2025.2607687

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, *160*, 108386. https://doi.org/10.1016/j.chb.2024.108386

Stenhouse, L. (1987). Artistry and Teaching: The Teacher as Focus of Research and Development. In *Media, Knowledge and Power*. Routledge.

Stöhr, C., Stathakarou, N., Mueller, F., Nifakos, S., & McGrath, C. (2019). Videos as learning objects in MOOCs: A study of specialist and non-specialist participants' video activity in MOOCs. *British Journal of Educational Technology*, *50*(1), 166–176. https://doi.org/10.1111/bjet.12623

Suchman, L. (Ed.). (2006). Interactive Artifacts. In *Human-Machine Reconfigurations: Plans and Situated Actions* (2nd ed., pp. 33–50). Cambridge University Press. https://doi.org/10.1017/CBO9780511808418.006

SUHF [Association of Swedish Higher Education Institutions]. (2016). *Recommendations for teaching qualifications in higher education* (REK 2016:1). Sveriges universitets och högskoleförbund/ The Association of Swedish Higher Education. https://suhf.se/publikationer/rekommendationer/

Sukhera, J. (2022). Narrative Reviews: Flexible, Rigorous, and Practical. *Journal of Graduate Medical Education*, *14*(4), 414–417. https://doi.org/10.4300/JGME-D-22-00480.1

Sung, C., Dhamecha, T. I., & Mukhi, N. (2019). Improving Short Answer Grading Using Transformer-Based Pre-training. In S. Isotani, E. Millán, A. Ogan, P. Hastings, B. McLaren, & R. Luckin (Eds.), *Artificial Intelligence in Education* (pp. 469–481). Springer International Publishing. https://doi.org/10.1007/978-3-030-23204-7_39

Suzen, N., Gorban, A., Levesley, J., & Mirkes, E. (2020). Automatic Short Answer Grading and Feedback Using Text Mining Methods. *Procedia Computer Science*, *169*, 726–743. https://doi.org/10.1016/j.procs.2020.02.171

Swedish Research Council. (2024). *Good Research Practice* (Text Nos. 978-91-89845-30–5). https://www.vr.se/english/analysis/reports/our-reports/2025-07-03-good-research-practice-2024.html

Thoppilan, R., De Freitas, D., Hall, J., Shazeer, N., Kulshreshtha, A., Cheng, H.-T., Jin, A., Bos, T., Baker, L., Du, Y., Li, Y., Lee, H., Zheng, H. S., Ghafouri, A., Menegali, M., Huang, Y., Krikun, M., Lepikhin, D., Qin, J., … Le, Q. (2022). *LaMDA: Language Models for Dialog Applications* (arXiv:2201.08239). arXiv. https://doi.org/10.48550/arXiv.2201.08239

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, *10*(1), 15. https://doi.org/10.1186/s40561-023-00237-x

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, *LIX*(236), 433–460. https://doi.org/10.1093/mind/LIX.236.433

UNESCO. (2022). *Recommendation on the Ethics of Artificial Intelligence*. (UNESCO Digital Library). https://unesdoc.unesco.org/ark:/48223/pf0000381137

Universitetskanslersämbetet. (2024, April 4). *AI skapar nya kompetensbehov—Högskolorna behöver anpassa sitt utbildningsutbud—Universitetskanslersämbetet* [Text]. https://www.uka.se/om-oss/nyheter/nyhetsartiklar/2024-04-04-ai-skapar-nya-kompetensbehov---hogskolorna-behover-anpassa-sitt-utbildningsutbud

Vallor, S. (2016). *Technology and the Virtues: A Philosophical Guide to a Future Worth Wanting*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190498511.001.0001

Vallor, S. (2024). The AI Mirror. In S. Vallor (Ed.), *The AI Mirror: How to Reclaim Our Humanity in an Age of Machine Thinking* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780197759066.003.0002

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems*, *30*. https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Verbeek, P.-P. (2005). *What things do: Philosophical reflections on technology, agency, and design*. Penn State Press.

Verbeek, P.-P. (2008). Cyborg Intentionality: Rethinking the Phenomenology of Human?Technology Relations. *Phenomenology and the Cognitive Sciences*, *7*(3), 387–395. https://doi.org/10.1007/s11097-008-9099-x

Verbeek, P.-P. (2011). *Moralizing Technology: Understanding and Designing the Morality of Things*.

Verdegem, P. (2021). *AI for Everyone?: Critical Perspectives*. University of Westminster Press. https://doi.org/10.16997/book55

Voordijk, H., & Vahdatikhaki, F. (2022). Virtual Reality learning environments and technological mediation in construction practice. *European Journal of Engineering Education*, *47*(2), 259–273. https://doi.org/10.1080/03043797.2020.1795085

Whittlestone, J., Nyrup, R., Alexandrova, A., Dihal, K., & Cave, S. (2019). *Ethical and societal implications of algorithms, data, and artificial intelligence: A roadmap for research*.

Winner, L. (1980). Do Artifacts Have Politics? *Daedalus*, *109*(1), 121–136.

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. F. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, *21*(1), 40. https://doi.org/10.1186/s41239-024-00468-z

Yang, Y., & Gulbahar, Y. (2025). Automatic Grading of Student Work Using Simulated Rubric-Based Data and GenAI Models. In J. Wilson, C. Ormerod, & M. Beiting Parrish (Eds.), *Proceedings of the Artificial Intelligence in Measurement and Education Conference (AIME-Con): Works in Progress* (pp. 34–39). National Council on Measurement in Education (NCME). https://aclanthology.org/2025.aimecon-wip.5/

Yuan, L., & Powell, S. J. (2013). *MOOCs and open education: Implications for higher education* [Report]. Cetis. https://www.cetis.org.uk/

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, X. (2022). *ChatGPT User Experience: Implications for Education* (SSRN Scholarly Paper No. 4312418). https://doi.org/10.2139/ssrn.4312418