

The Impact of Artificial Intelligence and Digital Transformation on Society: An Interdisciplinary Study of Human Behavior, Education, Governance, and Social Change

¹Muhammad Minhaj, ²Dr. Muhammad Habib, ³Faisal Rahman, ⁴Sara Nasir,
⁵Dr Madiha Hussain, ⁶Iram Shehzadi, ⁷Asad Ullah

¹PhD Student Business Management & Technology ALFA UNIVERSITY COLLEGE, MALAYSIA

Email: minhaj3@gmail.com

²Assistant Professor Department of Humanities and Social Sciences Bahria University Karachi

Email: muhammadhabib.bukc@bahria.edu.pk

³Department of IT Management Southwest Baptist university 4431 S Fremont Ave, Springfield, MO 65804 USA

faisal.rahman@sbuniv.edu

⁴Assistant Director (Community Development and Participation) Diamer Bhasha Dam Project WAPDA

Email id: eng.saranasir@gmail.com

⁵International Islamic University Islamabad,

madiha.hussain@iiu.edu.pk

⁶Visiting lecturer in communication and Media Studies at University of Sargodha

iramshahzadi235@gmail.com

⁷Department of Educational Research & Assessment, University of Okara, Okara, 56300, Pakistan

<https://orcid.org/0009-0001-1823-5743>

Email: asadullahsial786@gmail.com

HOW TO CITE: ABSTRACT:

Muhammad Minhaj, Muhammad Habib, Faisal Rahman, Sara Nasir, Madiha Hussain, Iram Shehzadi, Asad Ullah (2026). The Impact of Artificial Intelligence and Digital Transformation on Society: An Interdisciplinary Study of Human Behavior, Education, Governance, and Social Change. International Journal of Special Education, 41(15s), 1329-1363

Artificial Intelligence (AI) and digital transformation are really redefining the nature of today's societies and the behaviors of people, education systems, governance and other processes of social change. Domains have been studied individually in previous studies and there are only a few studies that examined the domains through a qualitative interdisciplinary perspective that critically examines interrelated social implications of these domains. This study aims to bridge this gap by examining the opportunities and challenges of innovation, inclusion, and efficiency that emerge as a result of AI-driven digital transformation, as well as the ethical, social, and governance issues that arise. The study is carried out from the interpretivist paradigm and qualitative research design which employs systematic document analysis of peer-reviewed journal articles, government policy documents, reports of international organizations and contemporary case studies. Following Braun and Clarke (2006) and Creswell and Poth (2018), the patterns, meanings and relationships identified in the data are done from

COPYRIGHT STATEMENT:

Copyright: © 2026 Authors.

Open access publication under the
terms and conditions
of the Creative Commons Attribution
(CC BY)
license (<http://creativecommons.org/licenses/by/4.0/>).

a variety of disciplinary perspectives via thematic analysis. The findings reveal five interrelated themes: Change in human behaviour, in digital identity; Change in AI-supported education and lifelong learning; opportunities and challenges of digital governance and of public service delivery; ethical issues of privacy, algorithmic bias, accountability and digital inequality; and AI as a catalyst for social innovation and institutional transformation (UNESCO, 2021; OECD, 2024). Theoretically, the study will provide contribution by formulating an integrated interdisciplinary framework that will provide explanations for the dynamic relationship between technological advances, institutional adaptation, and societal transformation. It offers a fact-based perspective for policymakers, teachers, tech innovators, and public servants seeking to promote responsible, inclusive, and people-centric governance of AI. The study provides a novel qualitative synthesis from the literature that integrates theory from various disciplines, such as sociology, education, psychology, governance and digital transformation, to shed light on the impact of AI on the path of social futures which are equitable and sustainable.

Keywords: artificial intelligence, digital transformation, qualitative research, thematic analysis, digital governance, education, social change

Introduction

The last 20 years have seen a dramatic technological transformation and Artificial Intelligence (AI), big data analytics, cloud computing, machine learning, robotics, the internet of things (IoT) and many more have become commonplace. All told, these innovations have made for a process that has been broadly termed "digital transformation," transforming the way people communicate, learn, work, govern and participate in society. This AI revolution is unlike any previous one, which led to shifts in the economic production and social organisation, but has now a unique capacity to automate thinking, to assist humans in decision making, to generate content and to influence social behavior at a global scale (Brynjolfsson & McAfee, 2014; Russell & Norvig, 2021). AI is not only used in research laboratories and industrial applications but has also infiltrated the daily lives of most people as digital platforms, education, healthcare or any financial service, public administration and social communication. The world has

experienced a shift in thinking regarding AI's potential to enhance service delivery, productivity, and innovation within the public sector over the last several years, as well as its implications for ethics, social and governance concerns. (OECD)

The impact of AI technologies has captured attention on a global scale, leading to numerous academic and policy debates. Unlike previous automation that was primarily manual, the new generation of AI systems are being used to perform tasks that involve human thought – language processing, problem-solving, prediction, suggestion and even decision-making. This shift has been further catalysed by the advent of generative AI applications, which can generate content in a manner similar to human creation, such as text, images, code, and multimedia. As a result, AI is changing the nature of the interaction between humans and technology, and important issues of autonomy, trust, identity, creativity, and social interaction have arisen (Floridi et al., 2018; Dwivedi et al., 2023). The use of AI in social systems has also

sparked discussions around algorithmic bias, misinformation, surveillance, privacy, accountability and digital inequality, calling for expanded interdisciplinary knowledge of the effects of technological change. The implications of AI on society go beyond its technological capabilities to its influence on human behavior, democratic institutions, and collective well-being, with recent international policy discussions focusing on these aspects. (UNESCO)

Digital transformation is an emblem of the current societies. It involves more than the use of digital technologies, it is also a reconfiguration of institutions, organizational processes, social relations, and cultural practices. The digital transformation is changing the way knowledge is created, shared and used in all industries and in the public sector. Digital infrastructures and AI-based systems are increasingly essential for helping governments to improve their services, make more efficient use of resources, combat fraud, and inform policy-making decisions with evidence. AI tools are being used in education to create intelligent tutoring systems, personalized learning environments, and AI-driven learning platforms to improve educational outcomes and accessibility. Meanwhile, companies are using AI technologies to streamline operations, improve customer experiences, and get strategic insights. The digital transformation is a social transformation, impacting on human values, behaviors and institutional arrangements, is reflected by the following. (OECD)

AI and digital transformation are reshaping people's behaviors, and that's a major part of the transformation of today's society. Digital platforms and artificial intelligence recommendation services have become more of an influence on individuals' information usage, opinion formation, social interaction and decision making. Social media and the patterns of communication and political behavior are driven by algorithms, and personalized content systems drive attention, preferences and behavior. Academic observers have reported

that cognitive functions are evolving due to the way people act in these environments, how they find out things, how they problem solve and how they form knowledge (Sunstein, 2017; Shoshana Zuboff, 2019). These technologies bring efficiencies and convenience, but can also result in an overabundance of information, algorithmic dependency, echo chambers and critical reflection deficits. Recent research also suggests adverse effects on human agency and decision-making as a result of the use of AI systems, which may have long-term implications for cognitive offloading and reliance on technology. (arXiv)

AI and digital transformation permeate today's society and the education sector itself has distinct implications. Adopting AI in education has created opportunities for individualised learning, adaptive testing, education administration efficiency, and accessibility to education. AI tools can personalize learning, identify weaknesses, and tailor learning experiences, improving learning. But there are ethical and pedagogical concerns that stem from these developments as well. In recent years, the topic of data privacy has become a prominent one in educational conversations, as has the question of algorithmic transparency, equity, academic integrity, and autonomy of learners. The study, conducted by the National Center for Supercomputing Applications (NCSA) at the University of Illinois, Urbana-Champaign, is a warning that the unrestricted use of AI technology could prove to be a danger to critical thinking, student agency, and educational inequities. Human-centred educational transformation thus places the emphasis on making sure that technology enhances rather than substitutes educational values. (DOI)

Governance systems and public institutions are in the midst of a transformation that extends beyond education, driven by AI. AI is being considered across the globe for its potential to enhance citizen engagement, service delivery, policy implementation and public administration. AI systems can be used to help governments manage vast amounts of data,

forecast social trends, boost efficiency, and better serve community needs. The governmental use of AI, however, has significant negative consequences for transparency, accountability, fairness, and legitimacy. Algorithmic decision making can exacerbate and propagate bias; opaque systems can eliminate democratic accountability and confidence in the system. With the increasing role of AI in society, the need for responsible and ethical management of AI, as well as protecting human rights and holding institutions accountable is becoming more evident than ever. The rise in the importance of trustworthy and human-centric AI signifies that there is a shared recognition that technological advancements need to stay within the framework of democratic principles and social values. (OECD)

The socio-transformative dynamics of the use of AI and digital technologies go beyond a specific sector or institution. Technological innovations impact on the labour market, economic organisation, cultural norms, political engagement and social relations. With the advent of automation and intelligent systems, employment is changing dramatically, presenting new opportunities and challenging conventional jobs. Digital technologies have opened new avenues for the sharing of knowledge, activism and collaboration; access to information and global connections. However, not equally distributed and it brings up the issue of digital divides and technological and socio-economic exclusion. The topic of access and different types of usefulness of artificial intelligence technologies to people, communities and countries is at the center of the discussions on sustainable development and social justice. International bodies have reiterated the need to highlight the importance of inclusive digital transformation strategies that not only ensure equal opportunities to participate in technology but also prevent its harmful impact. (AP News)

While there has been an increasing amount of literature on AI and digital transformation, there are still gaps in research. Much of the current

research literature addresses particular sectors, technologies, or policy concerns and, in some instances, a particular facet of these sectors, technologies or policy concerns – as in the case of research on applications of technologies in education that focuses on one technology facet or on educational uses, or on governance that focuses on a particular aspect of governance, or ethical issues that focuses on a particular aspect of ethics, or on organizational innovation that focuses on a particular aspect of organizational innovation. The studies offer important perspectives, but often neglect the multi-dimensional and interconnected nature of technological change. The influence of AI on human behavior, on education systems, on governance, and on social change in general, are widespread and varied, but relatively few studies have taken an interdisciplinary approach that can capture these aspects. Furthermore, the emphasis on quantitative technologies (and their effects) is often at the expense of the qualitative understanding, living with, and responding to technology led changes and transformations by people, institutions and communities. Thus, a need arises to consider the societal implications of AI in a more holistic and interpretive manner; this can only be done by means of qualitative research.

This study fills this gap by examining the impact of AI and digital transformation in society in an interdisciplinary qualitative manner. The study is not only about technological capability and economic effects, but also about the effects of this technology on human behavior, teaching, governance as well as on social change in general. The research aims to combine the views of sociology, education, psychology, public policy, technology studies, and governance studies to offer an overall perspective of modern digital transformation. The practice is particularly important as it is recognized that the impact and risks of technological systems are mediated by larger social, cultural and institutional systems.

This study aims to delve into and critically examine the social impact of AI and

digitalization in various areas. Research efforts will include examining the implications of AI for human behavior and social interaction, the impact of AI on educational systems and learning experiences, the implications of AI for governance and public administration, and the contribution of AI to the larger social change processes. Using qualitative analysis the study seeks to identify emerging themes, opportunities, challenges, and ethical issues in the transformation by AI.

The following research questions will guide in the achievement of the above objectives:

How is the use of AI impacting today's society's behavior and decision-making process, and how does it influence social interaction?

2. What are the opportunities and challenges of bringing the use of AI and digital transformation into the education environment?

3. What is happening to governance systems, public services and institutional accountability due to AI?

4. What are the larger-scale social shifts that are being created in the wake of AI and Digital transformation?

5. What kind of moral, social and political issues may arise from the use of AI in society?

This research may be considered an academic success and a public policy. Academically, the research builds on a growing interdisciplinary literature on AI and society, by bringing together various academic disciplines and highlighting the interdependencies of technological change. In addition, the research contributes to a qualitative understanding of AI as it explores the experiences and social interpretations of AI and institutional dynamics related to it that are not often discussed in research that focuses on quantitative analysis. The results can be used to guide policy-making and facilitating the governance of AI in a way that is human-centric, inclusive, and ethical. More nuanced insights into the societal impacts are required to guide the design of strategies by policymakers,

educators, technology developers and public institutions, in order to harness the societal potential benefits and mitigate the social risks of AI. The soaring importance of such research in contemporary governance and development frameworks is further emphasized by this trend of developing trustworthy, accountable and equitable AI in the international sphere. (UNESCO)

This article continues with the following sections. The literature review then takes into account the most recent theory and empirical research on AI, digital transformation, human behaviour, education, governance and social change. The methodology section describes the qualitative research design, philosophical assumptions, data sources and data analysis method employed in this study. The findings and discussion section outlines the main themes found in the analysis and critically draws out their meaning in a wider socio-cultural context. Lastly, it summarizes the main findings, suggestions for theoretical and practical contributions, limitations, and future studies and policies in the era of AI-driven digital transformation.

Literature Review

Introduction to theoretical background and interdisciplinary perspectives of AI and digital transformation

Artificial Intelligence (AI) and Digital Transformation are key elements of the Fourth Industrial Revolution, impacting and changing social, economic, educational and political life in every part of the world. Historically, technology has always had an impact on human society; but in this era, the rapid evolution of AI has resulted in unprecedented changes to the society in which humans live, with machines doing what was previously believed to be human-based tasks, including pattern recognition, learning, reasoning, decision-making, language generation, etc. (Russell & Norvig, 2021). AI has the potential to enhance, or even replace, cognitive processes, unlike previous technological shifts which were mostly aimed at automating manual

work. Therefore, researchers in different areas argue that the implications of AI are not restricted to merely being a technological improvement, but extending to being a socio-technological phenomenon with effects that are far from the principle of computational efficiency (Floridi & Cows, 2019; Dwivedi et al., 2023).

Digital transformation isn't unique: It's not simply a matter of digital technologies. A growing paradigm in the field of digital transformation today is that of a process: digital technologies are reshaping organizational structures, institutional practices, cultural norms and social relations (Vial, 2019). Digital transformation is not about replacing technology, it's about a continuous interplay of technology and people. This broader understanding would take into account the social, political, economic and cultural contexts in which technological systems are deployed and the impacts that they have on these contexts. As such, the relationship between AI and digital transformation cannot be separated and is being explored in these sectors, such as education, healthcare, public administration, finance, manufacturing and communication, where AI is one of the main technological accelerators towards digital transformation.

Theories of technological change exist which give insights into these developments and are relevant to them. Rogers' (2003) Diffusion of Innovations Theory is still one of the most popular models used to explain the process of diffusion of technological innovations in a society. Rogers' model of adoption has five elements; perceived relative advantage, compatibility, complexity, trialability, and observability. Even though the theory has been applied to predict the diffusion of variety of technologies, the recent scholars have believed that AI is not a single adoption process given its evolving nature after adoption through the machine learning process (Dwivedi et al., 2023). This dynamic ability adds new ethical, organizational and governance concerns to Rogers' original concept.

Likewise, Davis' (1989) Technology Acceptance Model (TAM) has had a strong impact on studies of individual technology acceptance of digital technologies. The model suggests that it is intention to adopt and perceived ease of use that will predict technological adoption. The TAM in education, organisations and healthcare has been confirmed by many empirical studies (Marangunić & Granić, 2021). Yet, some scholars have suggested that models of technology acceptance do not sufficiently explain why users accept AI due to the importance of the issues of algorithmic transparency, trust, explainability, fairness, accountability, and ethics (Shin, 2021). Today, however, the scope of AI research has shifted from the traditional theories of acceptance to broadly-based socio-technical approaches.

In the past few years, interdisciplinary research has thus focused on human-centered AI as a new conceptual approach. In contrast, human-centred AI suggests that the development of technologies should be based on human values, well-being, dignity, autonomy and democratic principles, not just efficiency and economic productivity (Floridi et al., 2018). Following this, other international organizations like the OECD, 2019, and UNESCO, 2021, have also encouraged the responsible governance of AI, aiming to innovate the use of AI while also being responsible. Such frameworks acknowledge that technological systems are by nature social systems, as algorithms are created, trained and used in human institutional contexts that are defined by cultural, political and economic values. This technological neutrality is being challenged in the field of AI scholarship, and is sometimes not accepted as neutral. In recent years, technological neutrality has been a topic of debate in AI scholarship.

The multi-disciplinary nature of the research field of AI reflects the multi-layered effects that AI has on society. Completely different underlying questions would be addressed by computer science, which would focus on algorithm development and technical aspects, and by sociology, which would focus on

changing patterns of social interaction and inequality; by psychology, which would focus on cognitive adaptation and behavioural responses; by education, which would focus on pedagogical transformation; by political science, which would focus on governance implications; and by philosophy, which would focus on ethical questions about autonomy, responsibility, and justice (Coeckelbergh, 2020). These disciplinary perspectives are not mutually opposing but are increasingly merging because of the impact of AI in many facets of modern society. The interdisciplinary approach is characteristic of the latest scholarship on AI, as is the notion that technological issues can only be solved from a wide variety of perspectives and cannot be solved through purely technical methods.

The socio-technical systems theory is a second, but significant, perspective on inter-disciplinary research on AI. The theory of socio-technical theory was first suggested by Trist and Bamforth (1951) and then extended by Bostrom and Heinen (1977) that there is a continuous interaction and co-evolution of technological systems and social systems. Technological outcomes, in this perspective, are determined by the organizational arrangements, institutional cultures, regulatory frameworks, and human behaviors as well as technological capacities. In more recent years, there has been a shift to socio-technical models for addressing the societal impacts of the same technology in various countries, organizations and communities (Rai et al., 2022). The results indicate that the key to successful implementation of AI is not solely technical, but also a comprehensive grasp of the institutional landscape.

Alongside the socio-technical perspectives, the Actor-Network Theory (ANT) of Latour (2005) offers another influential framework for examining a change facilitated with AI. ANT undermines the traditional binary opposition of humans/non-humans by providing a new way of understanding technologies as agents of social network formation. The concept of algorithms as “actors” – whose actions can be interpreted

as having an influence on the decisions, communication, ordering of information and institutional practices – is gaining in popularity in the field of AI research. The concept of an “actor”, an algorithm whose actions can be interpreted as having an effect on decision-making, communication, information ordering and institutional practices, is more common in the world of AI research. Critics have pointed out that ANT's over-agnostic stance on agency for the artefacts has been criticised (Fenwick & Edwards, 2016), yet it is suggested that it can help to account for the dynamics between people and smart artefacts. This shift is gaining importance as AI systems engage in decision-making tasks traditionally held to the human experts.

There has also been some philosophical debate on the relationship between humans and intelligent machines, as AI and digital transformation have become a part of life. Pessimistic Scholars suggest that AI is a tool for augmenting, rather than replacing, people and that it will be more useful for human enhancement than human replacement. AI can enable people to perform complicated tasks more efficiently, to develop new solutions and to make better decisions together with people, as stated by Brynjolfsson & McAfee (2014). In the augmentation view, there is a focus on the interaction between human and AI intelligence, and the belief that technological advancements increase possibilities for creativity, productivity and problem-solving.

More critical views, however, worry that the use of AI will begin to replace humans in terms of reliance on automation. In Zuboff's (2019) theory of surveillance capitalism, the digital technologies are completely appropriating the human experience, gathering vast amounts of data and forecasting human behaviour through algorithms. The perspective implies that AI can influence the dynamics between tech companies, governments, and users. The outlook implies that AI can change dynamics between tech companies, governments, and users. Likewise, O'Neil (2016) provides examples of algorithmic

systems that can drive discrimination by making discriminatory decisions in the context of the workplace, criminal justice, finance, health and education. The non-technological approaches highlight the fact that AI systems can exacerbate social inequalities, as they are generally based on biased data.

This is an important consideration, backed up by recent empirical research that has revealed that AI systems can tend to mirror the gender, race, ethnicity, socioeconomic status and geographic disparities of real-world populations (Mehrabi et al., 2021). Even though algorithmic fairness has become a new and important branch of research it is unclear whether fairness is precisely definable in different social contexts. Others have offered mathematical measures of fairness, others argue that fairness is not a technical component of optimization that is computable but rather a social and political and ethical question which should be decided by democratic debate (Barocas et al., 2023). This disagreement represents other interdisciplinary conflicts between the different approaches from computational ones, focused on measurable results, and social scientific ones, focused on contextual interpretation.

Another major discussion is the technological determinism versus social constructivism. Technological determinists argue that technology can influence social change by altering institutions, economies and culture without involvement from other factors. Social constructivists, however, argue that societies influence technological development by means of political decisions, cultural values, economic interests and institutional priorities (MacKenzie & Wajcman, 1999). Modern AI scholarship is largely accepting the holistic approach to the influence of technology and society on each other, avoiding simple, deterministic explanations. AI is influenced by and reflects our humanity, governance structures, market forces, and cultural values, and vice versa. This mutual view provides a wider perspective on digital transformation than one that just

considers the technology to be an historical force.

In this wider interdisciplinary conversation, there is a growing awareness of the importance of the concept of trust as a key factor in understanding the acceptance of AI by the public. Technical reliability is not the only aspect of trust that matters; it is also related to perceptions of transparency, fairness, accountability, explainability, and institutional legitimacy (Shin, 2021). Users can reject the use of technically correct AI systems if they are not clear or match the cultural values when making decisions. Therefore, explainable AI is a well-defined field of research to improve transparency without sacrificing predictive power. But there's a split in the research community about whether or not explainability is a good way to build trust among the public. There are some empirical studies that show positive relationships between transparency and trust, but others show that if there is too much technical information, they can either confuse users or give false information about the algorithms' reliability (Miller, 2019). The conflicting results indicate that trust is a socially constructed phenomenon, and not just a matter of the technological attributes.

While there has been significant theoretical development in the literature, there are still a number of important gaps. Firstly, there is much more research to be done that is focused in a disciplinary manner and the implications of AI are not well understood. Pedagogy innovations are an interest of educational researchers, whereas governance scholars focus on the regulation of the educational system, computer scientists focus on algorithmic performance and sociologists focus on inequality; comparatively little research has integrated these perspectives into a single analytical framework. Second, several studies examining the impact of AI focus exclusively on the economically developed world, especially North America, Western Europe and East Asia, and are relatively few in developing areas, each of which has different institutional arrangements, digital infrastructure,

and cultural norms and values (UNESCO, 2021). Last, no quantitative studies have explored the interpretation and experience of transformation that results from the use of AI, especially beyond the lens of technology adoption, economic productivity, or system performance, and fail to provide strong evidence of what happens when people, communities, teachers, or even policymakers and institutions use AI.

Such restrictions are particularly significant when AI is evolving from a standalone application to a social infrastructure. Therefore, an understanding of the wider implications of such a development would require a qualitative, interdisciplinary analysis that would be able to represent the complex interactions between technological innovation, institutional change, human behavior and social values. There are theoretical approaches that have been proposed to understand AI and all of them collectively demonstrate that it is not possible to understand AI with a purely technological analysis. Instead, its social meaning becomes prevalent through ongoing interactions of algorithms, institutions, policies, cultures, and human agency. The next section discusses the conceptual approach of AI and digital transformation on the human impact on behaviour, cognition and social interaction, which is a deeper understanding of the impact of technological innovation on the present society.

Knowledge of AI and Digital Education Transformation

The use of Artificial Intelligence (AI) in education is now one of the most widely studied aspects of today's digital transformation. The application of AI technologies in education is steadily increasing on the international stage, including the use of EDPs, automated grading, virtual/augmented learning and AI-generated content in the field of education in recent years. Over the past few years, AI in education has been a growing trend and many different applications have been introduced in educational institutions globally, such as: Education Data

Analytics, AI-powered assessment, Virtual and Augmented Learning Environments, and AI-driven content creation. AI systems have adaptive capabilities that can generate personalized learning experiences, predictive analytics, automated feedback, and intelligent tutoring and real-time learning support, which previous technology in education could not provide (Holmes et al., 2022). AI is not only changing the interactions between teachers and students and the way learning is acquired, but also the conception of knowledge and pedagogy, of the role of the learner, of autonomy, and of the teacher in a digital learning environment.

The recent studies in the field have largely recognized the potential of AI to enhance the quality of education through personalized learning and administrative efficiency. The intelligent tutoring systems, learning analytics platforms, and adaptive learning environments can assess student strengths and weaknesses and provide suggestions for learning pathways, feedback that would be difficult to provide in a traditional classroom setting (Zawacki-Richter et al., 2019). The studies show that such technologies can improve the effectiveness and efficiency of learning, stimulate student interest, and support inclusive learning for all learners in various learning styles (Luckin, 2018; Holmes et al., 2022). Beyond that, AI technologies have proven helpful in administrative tasks such as enrollment management, curriculum management, plagiarism detection, academic advising, and allocation of institutional resources, permitting educators to devote more time to educating and supporting students.

The introduction of generative AI has further fueled the shift of education and transformation, providing educators and students with some high-tech possibilities to create essays, coding, research notes, lesson plans, visual images, and interactive learning resources. In the realm of education, these LLM applications, especially ChatGPT, have swiftly infiltrated classrooms, classrooms, and beyond, providing novel opportunities for collaborative learning, creativity, language development, and

pedagogical support (Kasneci et al., 2023). Such technologies could become cognitive companions that complement critical thinking, problem-solving skills and access to educational resources, particularly when there is a lack of institutional support for students (Dwivedi et al., 2023). Last, AI should be seen as an educational tool for supporting, not supplanting, the human mind, and the teacher should be able to spend more time on teaching skills of thinking in a more abstract way.

While these are all encouraging strides, there are significant pedagogical and ethical issues that have been pointed out in the literature with regard to the integration of AI. One of the big issues up for discussion is that of the loss of critical thinking and independent learning. The fear is that relying too heavily on AI-generated content may lead to superficial learning, reduce students' intellectual involvement, and impact on their analytical and reflective abilities (Selwyn, 2019). The transition from knowledge building to knowledge consuming in learning processes might happen when students are using AI to help them craft assignments, abstract literature, and come up with solutions to complex problems. This reliance has spurred concerns about the true impact of AI on learning versus its efficiency.

The evidence based answers to this question are still unclear. There are also research papers that argue that the value of AI-assisted learning can be boosted if learning environments are carefully designed to allow reflection, collaboration and critical inquiry (Holmes et al., 2022). But others disagree that the widespread adoption of generative AI may reduce students' desire to engage in critical thinking and to accept algorithm-generated information (Kasneci et al., 2023). The contradictory findings suggest that the effects of AI on learning outcomes are less about the technology, and more about instructional design, teacher competencies, and the institutions' policies—along with the digital literacy of the learners. This has resulted in more and more thought by researchers and academics on the role of AI as an supplement to more

important human learning experiences rather than a replacement for them.

Academic honesty is also a huge issue which is discussed in the current literature of education. Generative AI's ability to create complex academic papers poses new questions about the nature of authorship, originality, and intellectual contribution. Given the potential for changed conceptions of independent learning, universities across the globe are rethinking assessment methods and the traditional, written assignment is no longer a valid measure of students' independent work (Cotton et al., 2024). Some scholars argue that the school should raise the standard of educational institutions and develop AI detection devices, while other scholars propose that the educational system should change by rethinking assessment methods, emphasizing the improvement of critical thinking, oral examination, group work, and problem-solving in practical scenarios rather than relying on memorization and repetition. It's a larger battle that's unfolding between visionary technology and tradition around pedagogical best practice.

Equity and access is also an important area of research. The power of AI to bring equality to the playing field and create an equitable learning environment is great – and with AI learning anywhere, to anyone, for anyone, the opportunities are endless. Intelligent translation systems, speech recognition technologies and adaptive learning platforms can be used to improve the inclusion of multilingual learners and students with disabilities in the education system (UNESCO, 2023). Yet, researchers at the same time warn of the potential for the digital infrastructure, the technological resources and the lack of AI literacy to accentuate educational inequalities, instead of narrowing them (van Dijk, 2020). There are often inequalities in the access to the internet, digital devices, and high-quality AI applications for students from economically disadvantaged communities. Therefore, technology, by itself, is insufficient to provide educational equity unless it is

complemented with investments in the digital infrastructure and inclusive educational policies.

Teachers' evolving professional identities are also an important area of enquiry. Based on most of the recent research, AI does not replace teachers but is rather used to complement and assist them in their task of teaching instead of the task of teaching. The majority of the latest studies show that AI does not take the place of teachers but rather provides support to them in the task of assisting them in teaching, rather than teaching. In the future, teachers will be seen more as facilitators for students to evaluate information created by AI, develop digital literacy, and understand ethical issues surrounding the use of technology. While many teachers report that their institutions are not supporting them to make the most of AI in their teaching, evidence does show that many teachers lack professional development and institutional support to effectively utilize AI in education (Zawacki-Richter et al., 2019). The mismatch of technology innovation and teachers readiness have been consistently identified in the literature, particularly in the developing countries where teachers' training is relatively less.

The theory of AI in education has also undergone a great deal of change. Early research tended to be based on theories of constructivist learning (knowledge construction by learners). More recently, connectivism is posited as a new paradigm in scholarship: Learning by making connections and relating to multiple sources of information across digital information networks (Siemens, 2005). In the context of an AI-enhanced environment, connectivist viewpoints come into particular focus as the learners are more often surrounded by smart systems that are generating, organizing, and recommending information. Critics argue that connectivism falls short in ethical, power and algorithmic implications in the digital learning space. To this end, contemporary educational researchers have begun to combine educational theories with ethical perspectives on AI to promote responsible technology use.

Overall, the literature indicates that the transformative possibilities of AI in education are well agreed upon, while there are also many challenges regarding the pedagogical, moral and institutional implications. Previous research tends to support the hypothesis that when applied to educational contexts that emphasize human-centred learning aims, such as critical thinking, transparency, and equity, AI can positively influence the learning outcomes. But there are conflicting views about the lasting impacts of relying on AI, on the matter of academic honesty and how to strike the right balance between technology and educational principles. Moreover, most of the current research has been based on the technologies used and the measurable learning outcomes, and there are relatively few studies that have used a qualitative approach to study the lived experiences of learners, the perspectives of educators, the institutional cultures, and the social ramifications. This is an imbalance that further underscores the ongoing demand for interpretive research that can explore the lived and understood experience of AI in a variety of educational settings.

AI in employment, work and economic transformation. The application of AI in jobs, work and economic change

The question of the link between Artificial Intelligence (AI) and employment has become one of the most debated topics in the research of digital transformation, which is an interdisciplinary topic. The relatedness between Artificial Intelligence (AI) and employment has become one of the most discussed topics in interdisciplinary research on digital transformation. Today, in the age of machine learning, robotics, natural language processing (NLP), and automation technologies, the way work is organized, managed and done has changed in both the public and private sectors. Historically, technology has driven changes to the labor market in large part by replacing human labor with machines, but in today's era, technology is increasingly automating complex cognitive functions, such as analysis,

communication, prediction and decision-making, as well (Brynjolfsson & McAfee, 2014). So, the researchers are now more inclined to view AI as a game-changer that will transform the way jobs are done, the skills needed, organizational practices, and the economy in general.

One of the prevailing views in the literature focuses on the productivity benefits to be gained from AI. The potential of AI to improve the efficiency and effectiveness of organizations, reduce administrative tasks, improve decision-making processes, and drive innovation and creativity through data analysis and insights has been highlighted in many studies (Bughin et al., 2018). AI-powered automation has been associated with various sectors that have experienced improvements in manufacturing, logistics, healthcare, finance, agriculture, customer service, and professional services, among others, and has helped to drive economic growth and competitiveness. AI is not about replacing human workers, but this is about moving them to jobs that involve more creativity, analysis, and interpersonal skills, and where human strengths still count. Proponents say that AI doesn't replace human workers but directs them toward more creative, analytical, and interpersonal tasks where human skills are still needed. (Brynjolfsson et al., 2022) This augmentation view sees AI as a technology that can augment, not supplant, human skills.

The empirical results back the optimistic interpretation to some extent. Previous research has demonstrated that, in most cases, AI assistance for professionals can boost productivity and performance by complementing human intelligence, not merely replacing it (Noy & Zhang, 2023). In industries such as software development, legal research, healthcare diagnostics, and scientific research, where decision making often requires complex judgments, AI serves as a valuable tool for professionals to consider when making a decision. The results question a deterministic view that automation of technology leads to widespread unemployment.

However, there is a large body of literature that comes with a more nuanced view of the labor market impact of AI. The influential study by Frey and Osborne (2017) set the tone for much research discussion on technological unemployment and predicted many jobs to be vulnerable. Subsequent research has disagreed with the magnitude of the automation threat, but there is agreement that AI has a disproportionate effect on repetitive, predictable, and rule-based jobs (Autor, 2015). Common areas targeted as areas of significant technological disruption include administrative support, routine accounting, manufacturing, transportation, customer service and data processing.

But, there is quite a bit of “noise” in the literature on whether AI will kill jobs or create jobs. Pessimists have been proven wrong in the past regarding the creation of new jobs, industries and business opportunities that offset the loss of jobs due to technological advancements (Brynjolfsson & McAfee, 2014). To support this, new positions are created in AI development, Cybersecurity, Data Science, Digital Ethics, Algorithm Auditing and Technology Governance emerge. Some, however, doubt the resemblance of the current AI technologies with those of the past – in particular, the fact that AI now increasingly performs higher-level cognitive functions as well as lower-level physical ones (Susskind & Susskind, 2022). As AI continues to progress in so many professions, the conventional concepts of the labor market adaption might need to be given a second thought.

The second is one of great importance, which is the changing nature of the skills needed in AI Economies. Recent studies have been continually stressing that future workplaces will increasingly prioritize skills outside of job-specific technical abilities, such as creativity, critical thinking, emotional intelligence, ethical reasoning, interdisciplinary collaboration, and lifelong learning (World Economic Forum, 2025). A digital literacy has therefore become an integral part of workers' skills that goes beyond

mere technical competency to cover knowledge of how to use AI tools, assessing AI-generated content, safeguarding digital privacy, and staying current with technological advancements. This has led to an urgent need for the rethinking of curriculum and skill development in educational institutions and workforce development programs, so that people can be prepared for a new job world increasingly dependent on AI.

Even though there is a general consensus that there is a need for reskilling, researchers have found there are still inequalities in workforce adaptation. In general, high-skilled workers will have better opportunities to become more productive and further their careers with AI, while lower-skilled workers will face a risk of being displaced, have stagnant wages, and be at risk of job insecurity (Acemoglu & Restrepo, 2020). This lack of economic benefits of technology further contributes to the broader socioeconomic inequalities and raises important questions of social justice, labour rights and inclusive economic development. Many of the scholars therefore suggested the coordinated enactment of public policies for workforce transition, lifelong learning, social security, and leveraging digital opportunities, in addition to market-based approaches, to address the challenges of AI-driven employment.

Similarly, organizational studies indicate that, regardless of the technology, the effectiveness of deployment of AI is more to do with the institutional culture, leadership, employee engagement and trust. There are common concerns among employees about surveillance, algorithm performance measurement, workplace monitoring, and lowering professional autonomy (Kellogg et al., 2020). The process of recruitment, employee performance evaluation, scheduling, and promotion is increasingly being influenced by algorithms, and new forms of algorithmic management have emerged that impact employer-employee relationships. Some feel their decision making efficiency and consistency have been enhanced and others feel the decisions made are not fair, the reasoning is not clear to them, or that they have less control

over the decisions made. The results confirm, more generally, the socio-technical views that the implementation of technology is still strongly dependent on organizational and human aspects.

In total, the literature shows that the impact of AI on work is not limited to aspects of automation. AI is changing the way people work, what they do, how they are managed in organisations, how they learn and how they engage in the economy. However, even the latest research provides conflicting results with regard to the long-term effects of technology, which vary by industry, occupation, institutional environment and nation. Also, much of the literature available is focused on economic forecasting and quantitative analysis of the labor market, while there is a dearth of qualitative analysis and foretelling of what the agent-driven transformation might mean for workers, employers, educators, and policymakers in real life. The interpretive aspects are important to consider when fostering a broader understanding of the broader implications of AI, particularly given the influence of digital transformation on labour markets around the world.

Digital Inequality, Inclusion and the Digital Divide

The fast-paced development of Artificial Intelligence (AI) and the Digital Transformation has drawn more academic focus to the topic of Digital Inequality and the continuing gap between the young and old. New opportunities for communication, education, healthcare, employment and civic engagement have emerged due to technological innovation, but not in an equal manner across societies. In recent years, contemporary researchers have started to pay special attention to the positive and negative effects of AI on SESI, depending on the availability, affordability, governance and inclusiveness of digital technologies (van Dijk, 2020). Thus, digital inequality has become one of the most prominent interdisciplinary issues

that have come into focus in the debates of AI-induced social change.

Research on the digital divide in this initial period tended to be concentrated on the points of access to computers and to the Internet (Norris, 2001). But modern research has acknowledged that digital inequality is more than just having or not having access to technology. Van Dijk (2020) suggests that it is important to recognize that digital inequality is a complex construct involving several interrelated aspects, such as access to digital infrastructure, digital skills, technology use patterns, and the ability to leverage digital engagement into positive social, educational and economic outcomes. It's a multi-dimensional strategy that has gained more traction as not everyone has access to the internet and it's not saying that everyone can access the benefits of AI technologies.

Empirical studies in recent years show that there are substantial country, regional, socioeconomic, and educational differences in the adoption of AI. Overall, the digital capacities of developed countries are more developed than those of developing countries, as are the funding sources for AI research, institutional capacity and the extent of AI regulation (UNESCO, 2023). This has led to a high prevalence of the application of AI for innovation in organizations and individuals in developed countries, particularly in the fields of education, healthcare and economic development. However, many developing countries are still grappling with problems linked to low internet connectivity, poor digital infrastructure, financial limitations, lack of technical skills, or lack of policy frameworks for the emerging technologies (World Bank, 2021). The disparities in the structures also contribute to the overall global inequities regarding access to and benefits from AI.

Socio-economic status continues to be an important factor in digital inequality. Many studies indicate that the impact of digitalization is much greater among those with high incomes, education and digital skills, and thus at a higher

risk of adopting AI technologies, and of using them in a meaningful way (van Dijk, 2020). Economically disadvantaged populations tend to experience marginalization in terms of access to digital devices, stable internet connections, technology literacy and AI-assisted educational materials. The gaps grow even bigger when it comes to having access to AI-powered learning platforms and digital tools – and don't. The gaps are even greater when it comes to access to AI-powered learning platforms and digital tools – and don't. AI can deepen existing inequities in educational outcomes in this context, if it is not widely available.

These structural inequities were further highlighted by the COVID-19 pandemic. In times of remote learning and digital service delivery, millions of learners had their learning disrupted due to poor technological infrastructure and lack of digital access (UNESCO, 2021). Many pupils could have continued with their schooling, but others could not due to cost, poor Internet connectivity or lack of technological infrastructure, aided by AI-powered pedagogical tools. Digital transformation, in turn, should be understood not as necessarily inclusive, unless it is supported by policy measures that will ensure access and participation among all.

Digital literacy has also become a key aspect of digital inclusion. It is becoming clear in current scholarship that there is a gap between access to technology and the skills needed to be competent with digital technologies. A digital literate person is able not only to operate with digital tools but to critically analyze and evaluate digital information, understand how AI systems work, possess an ethical awareness and awareness of online communication, data protection and informed decision making as well (Ng, 2012). These skills are more and more important for being able to engage in today's society. While people who are physically able to access digital technology are not necessarily digitally literate, if they are not digitally literate, they may be subject to misinformation,

algorithmic manipulation, privacy violations, and technological exclusion.

Geographical inequalities connected with digital transition are also highlighted in the literature. In rural areas, quality of Internet connectivity is often not adequate, investments in technology are lower, and the accessibility and availability of educational infrastructure are also lower than in urban areas (World Bank, 2021). These disparities also impact the ability to gain access to AI-driven health care, digital government services, online education, and employment. Although the timescale of recent breakthroughs in mobile technology have helped to reduce geographic inequalities, studies have long demonstrated that there are distinct urban-rural differences in development, urbanisation and poverty in developed and developing countries.

Another common topic in AI and digital transformation studies is the existence of gender gaps. Despite significant progress in the participation of women in digital technologies, gender gaps in AI education, AI-related technological employment, in leadership roles and in digital entrepreneurship (UNESCO, 2023) remain. The gender gap continues to be significant in the fields of computer science, engineering, and AI research, with women being underrepresented in these areas, which raises concerns about diversity in algorithm development and technological innovation. The lack of gender diversity in AI development teams can lead to algorithmic systems being more prone to reinforcing gender stereotypes and discriminatory outcomes, as training algorithmic systems with data that reflect historical social inequalities can create inequities and stereotypes within the system (Mehrabi et al., 2021). Thus, the pursuit of diversity in AI research and development is a crucial approach to enhance the fairness and inclusivity of algorithms.

The attention of scholars has also been increasingly focused on age differences. A few older adults face more difficulties in adjusting to the fast-changing digital technologies than

others due to either little digital experience, accessibility problems, or inadequate training (Czaja et al. 2019). The increasing integration of AI-powered digital services in everyday government, financial and health practices can lead to a loss of access to key government services for technologically disadvantaged groups. This phenomenon highlights that digital exclusion is not just an economic issue and involves issues of demographics and social disadvantage that need targeted policy responses.

While there is consensus on the need for digital inclusion, scholars have a variety of ideas about what works best to tackle digital inequality. Technological determinists often see infrastructure, Internet access and technological innovation as key solutions. Social scientists cautioned that efforts to overcome digital inequality are not limited to technological investment as exclusion is more rooted in the wider socioeconomic, educational, political and cultural inequities (Selwyn, 2019). From this point of view, integrated public policies should aim at two key areas: digital literacy programmes and social protection, as well as the reform of educational systems and the development of institutional capacities and resources, and distribution. This wider notion has gone on to receive growing backing in international policy approaches that champion a human-centred digital transformation approach.

Another weakness that was found in the literature is the geographical concentration. Although there has been a substantial amount of empirical research on digital inequality in North America, Western Europe and East Asia, there are comparatively few qualitative studies that investigate the adoption of AI in low-income countries or marginalised communities (UNESCO, 2023). Thus, current models and theories might not fully reflect the complexity of social, cultural and institutional factors in different geographical areas that shape digital transformation. However, more contextual, qualitative research is needed to grasp the

meaning of inequalities of diverse groups in everyday life related to AI.

Overall, the literature shows a consensus that digital inequality is one of the key challenges that accompany the transformation of the society through AI. The benefits of AI for social inclusion, education, economic development, and public service innovation can only be realized if AI is accessible, ensures digital literacy, involves institutional capacity, and governs inclusively. The philosophy of digital transformation is increasingly acknowledged in the existing scholarship, not only in terms of technological progress, but also in terms of the ability to achieve social justice, equal opportunity and sustainable human development.

Public policy, governance, and social change are explored in the context of Artificial Intelligence, with a view to critical analysis and research gaps

In recent years, the use of Artificial Intelligence (AI) in governance and public administration has completely revolutionized the government-citizen-institution dynamic. The adoption of AI technologies in public service delivery has become increasingly prevalent across various countries, aiming to boost administrative efficiency, optimize policy application, aid in evidence-based decision-making, and tackle complex social issues. In recent years, AI-based systems have been deployed in numerous sectors, including taxation, healthcare systems and administration, judicial systems, transportation systems, urban planning, immigration administration, disaster management, and public security (Sun & Medaglia, 2019). The developments show that AI has grown beyond being a technological innovation to an important governance tool that can affect institutional effectiveness and democratic processes. However, the significance of technology efficiency is increasingly highlighted, as the effectiveness of AI-supported public administration is ultimately rooted in principles of legitimacy, transparency,

accountability and public trust (Wirtz et al., 2020).

Many studies indicate that AI can help governments become more efficient by analyzing data, making predictions, and automating administrative tasks. The benefits of intelligent systems for governments include processing massive amounts of data, detecting emerging social trends, optimizing resource allocation and enhancing the quality of policy decisions (Engstrom et al., 2020). AI has proven to be a powerful tool for assisting quick government responses during public health crises, natural disasters, and economic downturns, by analyzing data in real-time and forecasting events. Likewise, AI-driven chatbots, automated licensing processes, and digital public services are becoming more prevalent in digital government initiatives to make governments more accessible, minimize bureaucracy, and increase citizen engagement. Such advancements can be used to substantiate claims that AI promotes more efficient, responsive, and evidence-based governance.

While all of these are possible advantages, researchers consistently warn that AI also presents many governance challenges. Algorithmic decision-making in public institutions is one of the most talked about issues. While traditional administrative processes are based on human judgment and legal reasoning, AI systems are often black-boxes, whose decision-making processes are complex and not easily known to citizens and politicians (Pasquale, 2015). This phenomenon is known as the "black box problem," and it has important implications for the issues of fairness of procedures, accountability of decisions in a democratic society, and the rights of citizens to challenge decisions made by the machine. A lack of transparency could lead people to feel less trust in government institutions when it comes to welfare distribution, criminal justice, job creation or allocation, immigration, or access to health care.

Algorithmic bias and discrimination is another significant issue. Many empirical studies have shown that AI systems can perpetuate or reinforce social inequalities; the algorithms that underlie machine learning can learn patterns from existing data sets that are imbued with social inequalities (O'Neil, 2016; Mehrabi et al., 2021). This means that AI-powered decision-making can lead to unintentionally discriminating against marginalized communities, defined by gender, ethnicity, socio-economic status, disability, and geographical location. The results cast doubt on the idea of technological neutrality and support the notion that AI systems are inherently influenced by their social, political, and cultural backgrounds. Hence, scholars suggest that while automation in decision-making is crucial, it must not replace human oversight, ethical scrutiny, or institutional accountability for effective governance of AI.

Privacy is also a key focus area in interdisciplinary AI scholarship. In the present day, AI systems make intensive use of vast amounts of data to make predictions, customize services, and fine-tune algorithmic efficiency. Being able to gather a lot of data makes technologies more effective, but it also raises issues of surveillance, unauthorized sharing of data, identity theft, and misuse of personal data (Zuboff, 2019). Organizations can easily gather extensive behavioural data and use it to anticipate consumer preferences, guide buying choices, and mold public opinion on digital platforms on a regular basis. Such acts foster the innovation of the commercial world and customization of services; they also raise substantial ethical issues on informed consent, informational autonomy and individual rights.

Their rise has raised scholarly discussion about the commodification of personal information, and the current state of surveillance capitalism has sparked a growing intensity in this debate. Zuboff (2019) says that digital technologies are more and more converting human experiences into economic goods by means of constant monitoring of behaviour and predictive

analytics. From this point of view, AI could not only improve the technological capabilities, but also transform the relationship between citizens, governments, and tech companies, as it will focus on an unprecedented concentration of information power in a few global digital platforms. A critical view of the data concentration points is that this could erode democratic institutions through the manipulation of public opinion, targeted political communication and mass surveillance. But those who support responsible data governance and proper regulation say it's possible for societies to be part of the AI revolution without compromising civil liberties and individual privacy. The opposing views offer a glimpse into the current conflicts between progress and democracy.

Trust is closely related to privacy, and is an emerging concept at the heart of AI research. Individuals' acceptance of AI technologies, sharing of personal information, acceptance of automated decision making, and involvement in digitally transformed societies are influenced by trust (Shin, 2021). Modern researchers focus on the importance of trust going beyond reliability to also involve ideas about fairness, transparency, accountability, explainability, institutional competence, and ethical responsibility. AI systems can be very accurate predicting the outcome, yet not well accepted by the public if they feel opaque or go against the societal values. Thus, trust is both technological and social in nature and is a result of interactions between algorithms, institutions, and citizens.

To address these challenges, the idea of explainable AI has arisen. Explainability aims to make algorithmic decision-making more comprehensible, by offering understandable explanations for the recommendations and predictions made by AI systems (Miller, 2019). The rationale for transparent systems is that they increase public trust by allowing users to assess algorithmic reasoning and find errors, if any. But empirical research on explainability is still limited. Some research shows that transparency and trust can go hand-in-hand, while some

shows that overly technical explanations can mislead or lead to false sense of security in information generated by algorithms. This conflicting research indicates that explainability is not enough to ensure trustworthy AI, and that there are additional institutional aspects, such as accountability, ethical regulation, and public engagement, that are needed.

Ethical issues are fast becoming a focal point in governance studies surrounding AI. International bodies, such as UNESCO (2021), the OECD (2019) and the European Commission (2020), promote human-centred AI that is grounded in human rights, fairness, transparency, accountability, privacy, non-maleficence, inclusiveness, and respect for human rights. These ethical frameworks acknowledge that the purpose of AI technologies is to improve human wellbeing, not simply to work more efficiently and/or to be more economically productive. Even as such, the question of how abstract ethical principles ought to be translated into concrete governance mechanisms that can manage the fast-changing technological systems is still debated. Some scientists support the idea of international laws that are legally binding, while others would like to see more flexible governance structures that foster innovation, but still compel ethical behavior. This uncertainty is indicative of the difficulty of governance in the context of technologies that change more quickly than legislation and regulation.

This is evident in the wider literature, where AI is emerging as a key driver of social transformation that goes beyond technological advancement. The uses of AI are expanding the range of its impacts on communication habits, cultural behavior, work relationships, educational environments, political engagement, health care, and social interactions in everyday life (Castells, 2010). The digital platforms have revolutionized the ways in which people obtain information, build social connections, engage in civic life, and develop identities. The algorithms used on social media platforms impact political communication, preferences and opinion

building, and AI-driven recommendation systems now drive cultural choices and consumer behavior. AI is thus driving significant shifts in personal attitudes and practices and social structures.

The potential for innovation, productivity, and social change through AI is being widely recognized across the field of research. As shown in the example above, there are many areas in which intelligent technologies can enhance human welfare, including medical diagnosis, scientific discovery, environmental monitoring, disaster management and educational accessibility (Topol, 2019). Some examples of the impact of AI include advancements in precision medicine, climate modeling, agricultural productivity, and accessibility services for people with disabilities. These are positive developments that once more foster positive expectations that this technology can be a contributor to sustainable development and quality of life if used wisely.

But the literature also points to high social risks of the AI-driven transformation. The reliance on intelligent technology could diminish human agency, heighten algorithmic control of everyday decisions, and lead to deterioration of human relationships through over-digitalization (Turkle, 2017). The potential of AI to enhance and diminish democratic societies further demonstrates this, with concerns raised by misinformation, deepfakes, algorithmically amplified polarization, and digital manipulation. Accordingly, scholars have increasingly pointed out that the impact of technology needs to be judged by its social implications, and not just by economic efficiency or computation.

Interdisciplinary studies on AI have increased significantly over the last decade, but there are still gaps in knowledge that need to be addressed. First, the current scholarship is very fragmented along disciplinary lines. The algorithmic approach, the technical performance, the productivity and the labour market results are frequently of interest to computer scientists, the learning environment is

of interest to educational researchers, and governance and regulation to political scientists. Very few research works consider these diverse views in a unified framework that captures the broad spectrum of the multidimensional implications of AI in society. However, in the context of the impact of technological change on individuals, institutions, and communities, interdisciplinary qualitative research is still important.

Second, most of the current literature still heavily centers on quantitative methods and models focused on technology uptake, economic metrics, algorithmic efficiency, or statistical analysis. These methods offer empirical insights, but often neglect the lived experience, contextual explanations, institutional cultures, ethical considerations, and social implications of AI incorporation. Qualitative research helps in understanding the social and cultural context and interpretation of educators, policymakers, professionals, citizens and organizations of the changes and challenges of AI-driven transformation. More such interpretive views are not as well represented in the literature.

Third, there is a lack of geographical representation in AI scholarship. Most of the seminal empirical research is conducted in North America, western Europe and East Asia, while developing countries are barely addressed despite the fast-growing digital transformation. The capabilities and structures of the institutions, regulatory frameworks, digital infrastructure, educational systems, and cultural values vary between different countries, indicating that AI could have different effects on society in different contexts. This would therefore enhance the global relevance of interdisciplinary AI research, if more focus is placed on underrepresented regions.

Lastly, although the ethics of AI is a growing area of research, current research on ethics, governance, education, employment, digital inequality and social change tends to focus on these aspects separately without exploring how

they relate to each other. There are some special reasons for knowing these relationships as technological change affects several aspects of society at once. Given these shortcomings, the present qualitative study aims to look at AI and digital transformation from a multidisciplinary perspective, exploring the nexus between AI and social change, and their intersections with governance, education and human behavior. The study seeks to synthesize and interpretatively analyze the opportunities, challenges, and societal implications of AI, in qualitative terms, to gain a more comprehensive view of the subject and to help develop responsible and human-centric approaches to digital transformation.

Overall, the literature has shown a wide consensus that the role of Artificial Intelligence and digital transformation are one of the most influential in today's society. However, there are still many questions about the impact of AI in the long term on democracy, employment, education, privacy, equality, and human autonomy. While opportunities for innovation and social progress are already great, so are the opportunities to highlight the importance of ethical governance, inclusive public policy, institutional accountability and responsible technological development, as existing scholarship has done. The ongoing interdisciplinary qualitative research is essential to address these unresolved debates, offering a chance to build a better understanding of the intricate relationships between AI, digital transformation, and society and to integrate various theoretical approaches. This form of approach serves as a conceptual basis for the current study and serves as a framework for the next methodological approach for the study of these interrelated societal transformations.

Research Methodology

Research Philosophy, Research Design, data sources, sampling strategy, data collection

The interpretivist approach to conducting research is employed to investigate multi-

dimensional implications of Artificial Intelligence (AI) and digital transformation on human behavior, education, governance and society in general. It fits the paradigm of interpretation because the research aims to explain the meanings, interpretations, and social realities of AI and not causal relationships or to test predetermined hypotheses. Interpretivism assumes that reality is created in different contexts as a social construction process by people's experiences, interactions and understandings, challenging the existence of multiple realities, depending on the context and the experiences of people (Creswell & Poth, 2018; Lincoln et al., 2018). The influence of AI in the modern society is multifaceted and has a wide spread impact on different sectors of society in different ways. Rather, a dialogue with ongoing scholarly studies, policy discussions, and social experiences is required to gain insight into the implications of AI for learning and teaching, governance, social relations, and human agency.

In an interpretivist approach to study, the emphasis is on subjective interpretation and a cultural or context-specific understanding of social phenomena, which are understood as varying over time, situated within specific cultural, political, historical and institutional contexts, and influenced by these contexts (Crotty, 1998). AI and digital transformation are not just technological innovations but socio-technical processes that give meaning as they interact with people, organisations, governments and societies. Hence, this philosophical perspective offers a holistic view of the opportunities and challenges that have been thought out by researchers, policies, teachers, and international organizations related to the adoption of AI. Therefore, the interpretivist paradigm is applicable in the philosophical approach to synthesize different perspectives and patterns that can be found in interdisciplinary literature.

The study adopts qualitative approach since it is employed to explore the rich context understanding of the complex phenomenon in

the social world and not a numerical measurement is required. Qualitative research is about the investigation of meaning, interpretation, experience and relationship in their natural settings, which aligns well with the study of the societal implications of AI and digital transformation (Denzin & Lincoln, 2018). This study is critical in its interpretation of the role of AI in different academic disciplines and policy areas, including education, governance, ethics, employment, digital inequality and social development, as opposed to the study that attempts to capture the statistical relationships and quantify the adoption of technology.

Justification of qualitative approach of design as it is a research topic which is interdisciplinary in nature. AI affects multiple areas of society simultaneously: institutional decision making process, the impact on education, public administration, labour market, ethical governance, etc., and also human behavior. It is not possible to fully capture these interrelated phenomena by a set of quantitative indicators since they are dependent on evolving social meanings, institutional practices, ethical debates and policy responses. Qualitative inquiry can then provide the opportunity to interpret the existing evidence from different disciplinary perspectives and make it easier to gain an understanding of the complex relationships. Others scholars also highlight the importance of using qualitative methods to study 'new technologies', particularly due to their ability to capture the social complexities, social interactions and evolving institutional responses which may go unnoticed in quantitative analysis (Tracy, 2020).

The method used in this study is qualitative document analysis which is a method used in studies to gain new meaning and understanding of existing documentary sources systematically (Bowen, 2009). Document analysis can be employed to study social phenomenon by analysing documents that are produced by humans that are not participating in the research process. Documentary evidence provides valuable information on debates in academic

research, policies, ethical principles, international recommendations and guidelines, institutional practices, and documented societal experiences in studies related to AI and digital transformation. Document analysis differs from the primary data collection methods like interviews or surveys as it is possible to integrate a wide and interdisciplinary body of knowledge that the researcher has obtained from various contexts and over an extended period of time.

The analysis of documents is especially suitable for the present study, given that the process of AI governance and the educational transformation, public policy, and digital ethics have created a wealth of documentary evidence generated by academic institutions, governments, international organizations, and policy agencies. The documents reflect a variety of views on technological innovation, societal transformation, ethical responsibility and regulation. As a whole, these resources can shed light on topics, conflicting viewpoints, emerging theoretical developments, and discussions in the field of AI scholarship. Furthermore, an open access to documents makes it possible to make cross-regional and cross-institutional comparisons that could help to grasp the societal impact of AI from a more multidimensional perspective.

More than one documentary sources are used to cover all aspects of the topic of the study. The primary sources of the information used are peer-reviewed journal articles in internationally published journals that are listed in the SSCI and Scopus indexes. The scholarly publications provide empirical and theoretical findings, methodological insights, and critical analysis of topics related to Artificial Intelligence, digital transformation, education, governance, public policy, ethics, employment, digital inequality, trust, privacy and social change. The top priority studies were the ones published in the past seven years (2019-2026) as they reflect the most recent trends and developments concerning generative AI, digital governance, and contemporary technological shifts. However, where these are needed, important key

foundational publications have been included for historical continuity and for theory.

Another source of documentary evidence is academic books. Foundational texts and theories surrounding AI, socio-technical systems, digital transformation and educational technology, governance, ethics and qualitative research engage with the recent empirical studies to add a theoretical perspective. Conceptualizing developments in technology, such as technology adoption, social construction of technology, the digital society and ethical governance, are still relevant to classics.

Government publications are another important source of information – with national governments adopting more and more policies, strategies and regulations on AI and as well as educational reform and governance policies to address new AI challenges. Policymakers find value in these documents for their opportunities, risks, regulatory agendas and implementation approaches they foresee for AI. Governmental reports can also be analysed in such a way that policy responses can be compared between different institutional and national settings.

Also important is the documentary data set from reports made by international organisations. Publications such as those from the UN (UNESCO, OECD, World Bank, World Economic Forum, European Commission and others) offer insights from an international perspective on responsible AI, digital inclusion, educational innovation, governance, sustainable development, and ethical principles. These organisations have been at the forefront of influencing the world's conversations on trustworthy AI, digital transformation and human-centred technological development. Consequently, they are good sources of authoritative information on international policy developments and new governance structures.

There is also valuable background on implementing AI in the education, public administration, health care and employment sectors of AI in digital governance, as well as AI policies and institutional guidelines. Such

documents often include suggestions for practice, regulations, ethics and implementation that build on academic research. Policy documents include theoretical aspect of the study, and practical governance as well.

Another source of qualitative evidence are case studies published in scholarly literature. Documented case studies provide context on how AI is affecting institutional practices and on societal outcomes in educational institutions, governmental organizations, healthcare systems, digital platforms and private organizations. This research looks at case studies not as stand-alone examples, but as a collection of evidence to find patterns of themes and common issues within AI-driven transformation.

In selecting a documentary evidence, the researcher employs purposive sampling method, which means that documentary evidence is obtained from relevant documents. In qualitative research, purposive sampling is a common technique used because the researcher can seek out information-rich sources who are capable of giving meaningful information about the phenomenon being studied (Patton, 2015). Purposive sampling is not about statistical representativeness, but about the documentation of documents of relevance, credibility and interest from a theoretical perspective. Because AI research is multi-disciplinary, purposive sampling allows for the inclusion of a variety of documentary sources from across a range of academic disciplines, institutional points of view, and geographical positions.

Several criteria were used to select the documentary resources. The first criterion for the documents was that they deal with one or more of the following: Public policy, social change, Trust, privacy, digital inequality, ethics, employment, governance, education, human behaviour, digital transformation, and Artificial Intelligence. Secondly, the papers were picked up from the peer reviewed journal articles published in the past five years from the reputable academic databases like SSCI or Scopus that generally have a higher

methodological quality and scholarly credibility. Third, the papers published over the past few years (2019-2026) were given priority, to promote the inclusion of recent developments related to generative AI and fast-changing digital technologies. The influential theoretical texts which have been published prior to 2019 were, however, kept if they offer important conceptual anchorages for modern research. Fourth, governmental information was chosen based on the institutional credibility and relevance to AI governance and digital transformation, as was international policy documents and organisational publication.

Inclusion/exclusion criteria

Document selection process was made transparent and consistent by using inclusion/exclusion criteria. Only documents explicitly dealing with Artificial Intelligence or digital transformation and the impact on society, education, governance, public policy, ethics, employment, human behaviour or interdisciplinary themes were included. The qualitative and quantitative empirical, theoretical, conceptual, policy and institutional reports were included if published. Only English-language publications were included because of limited resources and the high level of English used in the international AI scholarship.

Documents were also not included if they dealt with only the most technical aspects of algorithm development, but not the more general aspects of society. The abstracts of the conference, the editorial commentaries with little or no substantive analysis, the non-scholarly internet documents, the documents from promotional materials and the duplicate publications, which were not sufficiently methodologically and theoretically developed for inclusion in the analysis were not analyzed. Research that did not directly contribute to the research at a more advanced level on the societal impact of AI and digital transformation was also removed.

The data collection approach involved by systematic identification, retrieval, organization and preliminary review of relevant documentary information found in academic databases, institutional repositories, government websites and publications of international organizations. Relevant literature was found using a combination of keywords such as "Artificial Intelligence", "digital transformation", "AI governance", "education", "human behavior", "digital inequality", "ethics", "privacy", "trust", "employment", "public policy", and "social change". Documents retrieved were then sifted through using the inclusion and exclusion criteria established. Once documents were chosen, all eligible documents were extensively analysed to look for common concepts, theoretical arguments, methodological approaches, policy views and documented societal implications of AI and digital transformation.

Documentary materials collected were then sorted into major theme areas, corresponding to the study objectives, namely: human behaviour; education; governance; ethics; employment; and digital inequality, trust, and other social change themes. The subsequent qualitative content analysis, thematic analysis, and coding procedures were based on this systematic organization. With its structured approach, the research guarantees methodological coherence and allows to consider existing knowledge with a comprehensive and interdisciplinary approach to the impacts of AI on society and digital transformation.

Document Analysis, Qualitative Content Analysis, Thematic Analysis, Coding Process

In the present research qualitative document analysis technique is adopted as the main technique of data analysis to analyze and interpret systematically documentary evidence related to Artificial Intelligence (AI), digital transformation, human behavior, education, governance, public policy, ethics, employment and other aspects of social change. Document

Analysis is one of the effective ways in qualitative research that can be used for analyzing social phenomena with the existing documents, which contain plenty of contextual, conceptual and empirical information (Bowen, 2009). Qualitative document analysis is not just a summary of what has already been written, but aims at developing new interpretations by analysing patterns, meanings, relationships and assumptions in the multiple sources of documents. This is particularly important for interdisciplinary studies, where it can be used to "weave" together knowledge from different perspectives of academic disciplines, policy and institutional.

The strength of document analysis in this research is that it provides a depth of understanding of the concept, implementation, debate and governance of AI in different societal contexts. The field of education, as well as sociology, psychology, political science, economics, information systems, ethics and public administration have received a considerable amount of scholarly attention with regard to Artificial Intelligence. There have also been several policy papers, strategic plans, ethical guidelines and governance reports on the use of AI for transformation from governments, international bodies and regulatory institutions. The analysis of these various types of documentary sources can help highlight the themes that emerge from these texts, the different points of view and the new theories that are developed as well as the debates that are still to be resolved that may not be discernible when each text is considered individually.

Document Analysis was systematically and iteratively practiced with the methodological rigour. Candidates carefully read each of the selected documents and understood the objectives, theories, methodology, principle results, and contribution of the document to the overall development of AI scholarship. During this first reading step, attention was paid to the identification of the key concepts associated with technological transformation, societal impacts, educational innovation, governance,

ethical responsibility, digital inclusion, employment and human behaviour. Through this familiarity process, the researcher was able to develop a general picture of the documentary evidence prior to detailed analysis.

Then each document was read multiple times to make a second pass at the document to better understand the explicit and implicit meanings. The analysis was not limited to the findings of the different publications but also centered on the interpretation of the various authors regarding the concept of AI, the perception of technological changes, the formulation of ethical issues and the interpretation of social changes. Specific focus was placed on finding similarities and differences across disciplinary perspectives as the research of AI often has different theoretical underpinnings in different academic disciplines. Pedagogy is a key concern for educational researchers, governance and regulation for political scientists, productivity and labour markets for economists and inequality and social change for sociologists. These perspectives helped to create a more comprehensive interdisciplinary understanding of the impact of AI on society.

The researcher was interpretive in the document analysis and his philosophical stance to the study. Interpretivist research recognises that the documentary evidence is socially constructed interpretations which are shaped by the historical, institutional, cultural and disciplinary contexts (Lincoln et al., 2018). The documents were thus not understood as representatives of the real, but as places of meaning; where people involved in research, policy-making and institutions make meaning of AI and digital transformation from them. This view allowed an examination of the assumptions, values, priorities and conceptual frameworks that have become integral to the discourse of modern AI, but would not lead to an uncritical acceptance of documentation as universally applicable and context free.

Documentary evidence was continually compared in various types of sources to ensure

analytical rigour. In addition to scholarly journal articles, the following resources were analyzed: Government reports, publications from international organizations, policy documents and documented case studies. As a methodological triangulation process was used, both theoretical and practical aspects of the policy were discussed, which helped to strengthen the credibility and thoroughness of the analysis (Patton, 2015). Themes which arose in multiple documents would allow for greater confidence in the uniformity of the results. At the same time some conflicting interpretations were explored in depth to determine the differences in interpretation and their contexts.

The study also took into account document analysis, which was used for the analysis and systematization of the documentation that was found. One of the most prevalent methods of analysing text materials used in qualitative analysis is qualitative content analysis which identifies, categorises and interprets meaningful patterns in qualitative data (Schreier, 2012). While quantitative content analysis is concerned with the number of times information occurs in a text, qualitative content analysis is concerned with understanding meaning, relationships and context of information in the text. This is very much in line with the interpretivist paradigm as it is an interpretation approach as opposed to measurement.

Qualitative content analysis was conducted in an interrelated process in several stages. Documentary materials were initially analysed to provide meaningful units of analysis that included concepts, arguments, statements, interpretations and policy recommendations directly pertinent to the research goals. These meaningful units were then organized into larger conceptual categories of common ideas and repeated patterns in the entirety of the documentary set. There was no pre-determined classification of categories but rather they were developed over time as the documents were reviewed. This inductive approach implied having the analytical categories fit the documents and not the other way around.

During the categorisation, much time was spent conveying the context meaning. No individual statements were analysed separately, but in the constellation of their respective disciplines, institutions and societies. For example, when considering algorithmic fairness, the conversation in computer science is conducted differently than in politics or in the research of education because in fact, there are several theoretical foundations of fairness. The use of contextual sensitivity all through the analysis secured interpretive validity and minimised the dangers of oversimplifying interdisciplinary debates.

The qualitative content analysis also helped to carry out a cross situational analysis geographically and institutionally. The documents from various countries and organizations reflected the conflicting priorities in AI governance, educational changes, technological progress, and ethical considerations. These differences helped to pinpoint common concerns at the global level and specific interpretations of the change AI brings. The comparative analysis helped to deepen the understanding of the experience and regulation of AI in different institutional contexts.

The principle analysis technique of thematic analysis was conducted to identify, code and analyze themes from the documentary evidence, following the qualitative content analysis. Thematic analysis is one of the most flexible and rigorous qualitative research methods that can enable systematic analysis of patterns of meaning in complex qualitative data sets (Braun & Clarke, 2006; Braun & Clarke, 2021). The varied documentary sources were effectively brought together and synthesised by thematically analysing the evidence, both from the different sources and across the different aspects of AI and digital transformation.

Thematic analysis was carried out using the six phases of the analytic framework proposed by Braun and Clarke (2006). This initial stage consisted of frequent familiarisation work, in the

form of careful and repeated reading of all documents that were selected. This phase enabled them to immerse themselves into the data and facilitated the identification of ideas, concepts and patterns which were repeated, relevant to the research objectives. At all stages analytical notes were kept for recording initial interpretations and potential connections between the different elements of the documentary evidence.

The next step was to create a set of initial codes which would capture meaningful segments of text that discuss the social aspects of AI. Coding focused on problems related to human behavior, educational change, government, ethics, jobs, digital inequity, privacy, trust, accountability, technological innovations and social change in general. The coding process was flexible throughout the analysis and interpretivist inquiry recognizes that meanings can change as the researchers become familiar with the qualitative data.

In the third phase, similar codes were aggregated into more general themes that occurred frequently throughout the documents. Thematic development not only identified the similarities of concepts, but also examined the conceptual relationships between the different facets of the process of AI-driven transformation. Examples include issues on digital literacy, accessibility in education, and technological inclusion, which helped to build the overall picture of digital inequality, and of transparency, explainability and accountability, which helped to build the overall picture of trustworthy AI governance. This process facilitated the combination of a variety of disciplinary viewpoints into themed clusters.

The fourth stage was to analyse the themes submitted by the candidates and its internal coherence and distinctiveness. For each theme, the material was considered against the documentary evidence; an attempt was made to determine how accurate it was and whether it included the diversity of opinions of the selected materials. Conceptual overlap was enhanced,

combined or reorganized in themes to strengthen the analytical clarity. This was a review process that was repeated and as such, a meaningful pattern emerged in the themes presented, rather than an arbitrary classification.

Themes were clearly identified and developed in the fifth phase by making regular comparisons between the document sources. Each of the themes was discussed in the context of the overall theory, as well as its role in understanding the social dimension of AI and digital transformation. Themes were not considered as separate categories but rather the interrelatedness between different aspects of technological transformation was identified, revealing the connections between education, governance, ethics, employment, social behaviour and how they influence each other within the broader aspects of digital change.

The third step was to synthesize the themes into an overall analysis that answered the research question(s) of the study. This was not a description of content but an analysis of relationships between themes, a comparison of contrasting points of view in the literature, discussion of convergence of theory and a discussion of areas of disagreement and gaps in existing research. The thematic framework which emerged was used for discussion and findings in the next chapter.

Coding was a part of the qualitative content analysis and thematic analysis processes. The coding process was carried out manually so that there was a continuous interaction between the documentary material and the interpretative understanding of the material could be achieved. Manual coding allowed the researcher to analyze the nuanced conceptual connections, differences between disciplines and variations in context that might have been missed using automated coding. As a result of first round open coding, concepts identified were very general in nature and directly came from the documentary record. Codes were then iterated and compared as analysis was done to form higher-level conceptual categories that were related.

Throughout the coding process, the analysis was flexible because qualitative investigation acknowledges that interpretation is an ongoing process that is iterative with the data (Saldaña, 2021). Newly identified concepts were constantly related to previously coded materials and existing categories were refined and expanded, forming more elaborate thematic structures. Coding activities were used in conjunction with reflexive memo writing, which was used to record analytical decisions, ideas, new interpretations, and connections between themes. These analytical memos added methodological transparency by offering a systematic record of how interpretations have developed.

The qualitative document analysis, qualitative content analysis, thematic analysis and systematic coding were all combined to create a strong methodological framework for the analysis of the complex social consequences of AI and digital transformation. The analytical tools complement each other and enable an in-depth analysis of interdisciplinary data from documents, while maintaining the interpretivist approach in the philosophical foundation of the study. The analysis aims to produce a comprehensive understanding of the impact of AI on human behaviour, education, governance, and social change, through systematic comparison, iterative coding and critical thematic development. The procedures used to guarantee the trustworthiness, ethical integrity, methodological rigor, and overall credibility of the research findings are discussed in the following section.

The trustworthiness, ethical issues, methodological restrictions, and methodological quality of the study

Trustworthiness of qualitative research is crucial in establishing the credibility and the academic rigour of research findings. Qualitative research often assesses quality in terms of statistical reliability and validity but, from a qualitative perspective, the criteria are trustworthiness, which is meant to show that the interpretations

are believable, consistent, transparent and supported by the evidence gathered (Lincoln & Guba, 1985). In keeping with the interpretivist paradigm followed in this study, the four criteria suggested by Lincoln and Guba (1985) were used to establish trustworthiness, which are: credibility, transferability, dependability, and confirmability. It is a general list of the characteristics of qualitative research and yet the criteria recognise the interpretative and contextual nature of social enquiries.

Credibility, as defined by Lincoln and Guba (1985) is the accuracy of the results in representing the meanings in the documentary evidence, as well as the convincing interpretation of the phenomenon of the study. To ensure the credibility of the study, a wide range and high quality of documentary sources were used including peer-reviewed journal articles, scholarly books, policy documents, government documents, published case studies and reports from international organisations. A methodological triangulation approach was adopted to ensure that multiple documentary sources could be employed to generate the opportunity for cross-disciplinary, institutional and policy comparisons of evidence (Patton, 2015). The project did not rely solely on one type of document but was also based on both theoretical literature and empirical research and policy evidence, thereby enhancing the confidence in the consistency of the findings.

This was further strengthened by multiple reading and analysis of the chosen documents. There were multiple iterations of document review to continually better the codes, categories and themes and to reduce the risk of superficial interpretation. There were consistent findings across sources regarding areas of agreement, contradiction and contextual variance. The interpretations throughout this analytical process were always supported by the documentary evidence and depended on no assumptions or isolated observations. In addition, throughout the analytical process the researcher recorded their changing interpretations and analytical

choices in reflexive memo writing, adding to the credibility of the writing.

The second is transferability related to the extent to which the findings of the research can be transferred to contexts beyond the specific context studied (Lincoln & Guba, 1985). One of the ways to describe qualitative research is to say that it is not about statistical generalizations, but about thick descriptive understanding in the context of a researcher's findings that can be evaluated by a reader in similar contexts. In this study the transferability was enhanced by detailed description of the research design, philosophical assumptions, selection of documents and analysis and the development of themes. By documenting the methodological process efficiently, the results of the study can be assessed for the application of the results in other educational systems, governance structures, institutional contexts and geographical areas.

Documentary evidence of the international perspectives was also included, which helped to enhance the transferability. All the works have been chosen from studies carried out in various countries and institutional settings, and from reports of international organizations such as UNESCO, OECD, World Bank, European Commission, or the World Economic Forum. The wide documentary base allowed a multi-dimensional analysis of AI and digital transformation in different social, political and economic settings, providing a more transferable analysis to other settings than any one country or discipline. The study recognizes, however, that interpretations are affected by the context and need to be considered in the broader setting in which the use of AI technologies occurs and is regulated.

Dependability will be the consistency and transparency of the research process and the research procedure must be methodical, logical, and well documented, and the results of the research must be clearly documented and understandable (Lincoln & Guba, 1985). To ensure reliability, a research design was created,

starting data analysis, and each stage of the research process, including the selection of documents, inclusion and exclusion of documents, the coding process, the thematic analysis and interpretation of the findings, was documented before the data analysis was begun. The method of data collection and data analysis of this study was done methodologically that is then supported by systematic qualitative document analysis, qualitative content analysis and thematic analysis.

The methodological decisions, coding revisions, analytical memos and development of themes were recorded in a transparent manner during the research process. This documentation will help future researchers understand the analysis that took place during this study and will give transparency of how interpretations were developed from the documentary evidence. Besides, the application of the qualitative analysis method, including thematic analysis method recommended by Braun and Clarke (2006, 2021), which is a qualitative analysis method that has been formed and widely used, contributes to the trustworthiness of the qualitative analysis method and ensures methodological consistency in the research.

Confirmability is the fourth criterion which highlights the need to make sure that the findings of the research stem from the documentary evidence, and not from the researcher's personal beliefs, assumptions, or preferences (Lincoln & Guba, 1985). In interpretivist research, total objectivity can not be achieved as researchers are involved in the interpretive process. For qualitative researchers, however, reflexivity is an awareness that they are active participants in the process of knowledge construction, and that the interpretations are based on the empirical evidence.

The researcher has been reflexively aware and has critically questioned his own assumptions in an attempt to become confirmable, to position himself in the road of new interpretations that developed in the documentary materials. A large volume of documentary evidence was used and

any conflicting evidence found in literature was not ignored in making decisions. The analysis intentionally took into account contrasting theories, differences in methodology, and differing interpretations of policy in order to show that there was no focus on studies that would support the pre-established findings. This 'balance' helps increase the objective part of the analysis and at the same time maintains the interpretivist view of knowledge as a social construct and contextually situated.

The use of the authentic scholarly literature and governmental publications, as well as international policy reports is extensive, this contributes to the confirmability as authors' interpretations are supported by authoritative documents. In-text citations in the proper APA 7th edition format throughout the study allow for clarity of the source of theoretical ideas and empirical data used and allow the reader to judge the supporting evidence on their own.

Whichever it is, primary or secondary data, ethics is also quite an important aspect of qualitative research. Ethical responsibility is essential in the present study as it is based on the interpretation and synthesis of published intellectual work, so that there is no direct interaction with human. Therefore, the principles of international academic integrity (Orb et al., 2001) are followed in the study: transparency, honesty, responsible scholarship and principles of academic integrity.

Ethical concerns that emerge are how to appropriately represent existing scholarship. Every effort was made to accurately state authors' ideas without distortion, paraphrasing, and not misrepresenting their intent, and to avoid quoting, paraphrasing, or interpreting their original meanings. The research work carried out by scholars is critically evaluated and synthesized and may not be copied. All ideas, theories, data from published texts and policy proposals were properly cited using the correct APA 7th edition.

The documentary material selection was also based on ethical principles that focus on the

quality of the scholarly work, credibility and transparency in documentary material. Journal articles, academic books, official government reports and documents from international organizations and other well-known institutions were emphasized when selecting sources, with a particular focus on those that had been peer-reviewed. In order to ensure the integrity and reliability of the research evidence, the internet sources and opinion blogs, promotional materials and documents with inadequate academic evidence were excluded.

Confidentiality and participant protection concerns that are often raised in qualitative interview studies were not a concern in this study as it involved documents that were accessible to the public. The researcher, however, had a moral obligation to respectfully interpret all the documentary data, and not to offer biased or misleading interpretations of the state of current scholarship. Another goal of the study was to provide a balanced discussion by recognising the opportunities and challenges of AI and digital transformation instead of pushing particular ideological agendas.

Although the research design was methodologically sound, there are some limitations that should be noted. First, the study is based on only secondary, documentary data and not primary qualitative data from teachers, policy makers, technology creators, organizational leaders or lay people. Documentary analysis can provide a comprehensive interdisciplinary knowledge of the issue and direct interviews or focus groups provide other information related to the issue, including personal experiences and perceptions, and the interpretation of AI implementation in the context. However, future study might be improved by also incorporating a combination of documentary analysis and interviews, observation or a case study approach.

Second, the documentary sources showcase a range of academic disciplines and perspectives, in the majority of cases economically developed countries and particularly North America,

Western Europe and East Asia. Therefore, the literature currently available may not necessarily accurately represent the views of developing countries, with diversity in technological infrastructure and governance capacity, education and digital transformation processes. More research from underrepresented regions would help to build a more balanced understanding of the implications of AI on society from a global perspective.

Third, technology is developing too fast for there to be a natural methodological problem. AI is constantly developing with new instances of machine learning, generative AI, robotics and digital platforms. In this sense, published research, policy statements and institutional practices are constantly changing and in some instances, changes may need to be made to reflect additional technological advances. In order to overcome this deficiency, a selection of recent literature has been selected which ranges in date from 2019 to 2026, with some key theoretical pieces of literature that still have a conceptual relevance that is relevant to this present study.

A second limitation is language - The documentary analysis only consisted of documents published in English. While the majority of international AI scholarship is in English, there may be a lack of access to region-specific views and culturally diverse interpretations of digital transformation if other languages are excluded. Comparative studies in the future which add multilingual documentary analyses can offer more general international perspectives.

Finally, thematic interpretation is always a matter of interpretation which will always require the judgment of the researcher, as is the case in all interpretivist qualitative research. Systematic coding processes, reflexive analysis, methodological triangulation and abundant documentation were used to enhance the trustworthiness, but different thematic interpretations could be reached by other researchers who used the same documentary

evidence. This is not a fault, but part of the interpretive character of the qualitative method, as long as the analytical process is clear and based on evidence.

Overall, the methodology adopted in this study is appropriate and adequate for the analysis of the societal impact of AI and digital transformation. The research design adopted is qualitative with the methodology based on the interpretivist philosophical approach and purposive document selection, qualitative document analysis, qualitative content analysis, thematic analysis, and systematic coding to support the research aims. Furthermore, the application of current criteria of trustworthiness, ethical guidelines, open analytical methods and recognition of methodological limitations further increase the credibility and research rigor of the study. The thorough methodological approach enables a detailed examination of the impact of AI and digital transformation on human behaviour, education, governance, public policy and social change in general, providing a solid ground for the presentation and interpretation of the results of the study in the next chapter.

Findings and Discussion

The five themes that emerged from the qualitative thematic analysis of the selected articles from peer-reviewed journals, policy documents, reports from international organizations and selected case studies represent the social implications of artificial intelligence (AI) and digitalization. Findings indicate that AI is not merely a technological advancement, but a complex phenomenon that impacts various levels of practice, institutional practice, governance frameworks, and social structures. The themes build on each other and emphasize the interdisciplinary nature of the opportunities and challenges arising from AI, all of which are ethical, regulatory and social in nature. The themes that are presented below are the broad patterns found in the analysis.

Theme:1 Learning outcomes for Theme:1: To investigate how AI and behavioural change can affect human behaviour

This theme is about how AI affects people, interactions and communication. It delves into its influence on digital identities, consumption of information, interactions online, trust in AI systems, decision making processes, and mental health. It also shows the influence of AI platforms, the Recommender system, and the Generative AI on an individual's learning, working, communication, and socializing activities. The advantages are discussed, such as increased productivity and access to information, but also the dangers of the use of AI, misinformation, lack of control and new social interactions.

Theme 2: Artificial Intelligence and its impact on education, the world of work and digital skills

In this theme, students will examine the changes AI will bring to the education system and the future of work. It considers how AI can be applied to individualization of learning, intelligent tutoring, assessment, digital literacy, lifelong learning and workforce development. The shift of work patterns, new skills requirements, automation trends in the workplace and growing demand for digital skills will also be explored during the study. Apart from these opportunities, the topic of academic integrity, access to educational technologies, its impact on the labor market and the need for lifelong reskilling and upskilling in the digital economy are discussed.

Theme3: Master's programme in AI has a third theme namely: Theme3:AI and Governance, Public Administration and Policymaking

The theme is on the growing use of AI in the governance and decision-making process in the public sector. It addresses the application of AI in public service delivery, digital government, policy making, public service resource management and evidence-based public service

decision making. The results delve into the contribution of AI to administrative efficiency, transparency and innovation, algorithmic accountability, public trust and confidence, regulatory oversight, explainability, and democratic governance. The discussion emphasises the importance of setting up technologically innovative and fair, accountable and legitimate governance structures.

Theme4:Ethical Dilemmas, privacy and digital inclusion in the age of AI

This theme will examine the ethical and social implications of the ever-expanding use of AI. There are common issues regarding algorithmic bias, discrimination, privacy protection, governance of data, cyber security, surveillance, transparency and accountability. It also examines the issue of digital inequality related to the distribution of opportunities related to AI in various groups and geographical areas. The discussion highlights the importance of responsible AI use, safeguarding against violations of human rights, promoting ethical governance, and inclusive digital policies to ensure that technological advancements positively impact society without furthering social marginalization and unforeseen effects.

Theme 5, the socio-cultural aspects of the role of AI in society and sustainable digital transformation will be explored

The last theme thematically brings together the results of the previous themes, looking at how AI can fuel more general societal and institutional change. It covers the impact of AI on innovation, economic development, institutional adaptation and sustainable development, and the impact that AI has on the relationships between technology, society and governance. The analysis reveals that digital transformation can be realized best when supported by ethical leadership, participatory governance, digital literacy, and a joint policy formulation. This theme builds on the previous themes, offering a holistic perspective on the impact of AI on social transformation, and highlights the importance of aligning technology

with human values, social responsibility and sustainable development in the long-term success of digital transformation.

Conclusion

This qualitative research explores the socio-cultural challenges and implications of Artificial Intelligence (AI) and digital transformation in an interdisciplinary manner, taking a human-centric view on how AI and digital transformation affect human action, education, governance and the social change as a whole. The study not only looked at AI as a new technology but as a socio-technical phenomenon that is impacting institutions, human relationships, the public policies and the way people live. We analyzed existing research and scholarly literature, policy documents, government reports, publications of international organizations, and documented case studies using a qualitative document analysis, qualitative content analysis and thematic analysis, which showed that while AI offers opportunities to innovate, become more efficient, and advance progress towards a better society, it also raises complex ethical, institutional, and governance questions. The results show that the social meanings, institutional consequences and human implications of AI need to be understood, even beyond just the technological performance, in various cultural and policy contexts.

The qualitative analysis results show that AI is not only a game-changer for certain parts of society, but also for certain parts of society. The recurrent patterns in human behaviour, decision making, communication, education, governance, employment and delivery of public services as revealed in the literature reviewed confirmed that AI is affecting them all regularly. Intelligent technologies are now helping with knowledge production, repetitive work automation, personalized learning, data governance, and optimizing organizational productivity. In parallel, there are growing concerns about algorithmic bias, digital inequality, privacy, misinformation, surveillance, transparency and trust, the study says. The results suggest that

technology should not be assessed only in terms of its productivity and/or computational power, but also its contribution to improve social inclusion, ethical responsibility, democratic accountability and human well-being. The findings align with the human-centred AI approaches and reflect the importance of technology developments being aligned to social values and fundamental human rights as stressed by Floridi & Cowls (2019) and UNESCO (2021).

The interdisciplinary approach of understanding the concept of AI and the digital transformation is one of the most valuable aspects of this study. There was much current research on education and/or governance and/or ethics, but not on technological innovation. The study brings these three areas together in one qualitative frame, revealing the integration of the three societal impacts of AI and the lack to capture them individually from a disciplinary perspective. Redeveloping education affects workforce readiness, the governance structure can influence public trust in AI, ethical standards influence the use of technology, and digital inequality has the potential to impact educational, economic and civic opportunities. This intertwined view helps create a more holistic view of the impact of AI on people's lives, institutions, and social processes. The study thus reinforces the case being made in the academic literature for the need for interdisciplinary research in AI, which can tackle the multifaceted nature of the technology, society and governance (Dwivedi et al., 2023).

Theoretically, the study contributes to the growing literature in the qualitative field of scholarship in that it demonstrates the application of interpretivist inquiry to gain insight into the development of new digital technologies. On-going research in AI continues to be quantitative analysis of the uptake of technological innovations, algorithm efficiency, productivity or economic outcomes. Although these approaches provide valuable empirical lessons, they do not account for social meanings, institutional contexts, ethical

challenges and human experiences in the context of AI-driven transformation. This study demonstrates that AI is a tool, an artifact, a socially-constructed phenomenon and is not simply a technical system. The findings therefore resonate with the socio-technical theories, human-centred AI notions and interpretivist orientations that have taken into account the social aspect of technological innovation but have not viewed it as a separate driver of change (Vial, 2019; Coeckelbergh, 2020).

The education, policy makers, technology developers, organizational leaders, and public institutions are other groups with major implications from the study. The findings suggest that AI should be used to support the learning and teaching process, creative thinking and critical analysis, and not be considered as a replacement for meaningful human interaction or independent intellectual activity in the field of education. Education policies should prioritize emphasizing the need to be digitally literate and to use AI tools ethically, as well as promoting teacher professional development, and implementing assessment strategies that encourage critical thinking and the creation of AI-generated or automated content. Likewise, companies implementing AI technologies should focus on training employees, transparency, responsible innovation, and a continuous human oversight to guarantee that intelligent systems enhance people's capabilities instead of replace them.

The results show how important it is to establish strong governance frameworks in governments and public institutions to face the challenge of technological innovation and to adhere to democratic values such as accountability, transparency, and public trust. Using AI tools in the decision-making of public administration should not lack in any measure of inadequate legal protection, ethical guidelines, and human oversight to minimize algorithmic bias, and protect the right to the citizens. States and governments must also invest in accessible digital infrastructure, access to digital technology and digital literacy initiatives to mitigate digital

inequalities and make the benefits of AI more accessible to different socio-economic groups and regions. Intergovernmental governance approaches are also important for the majority of the ethical and regulatory challenges associated with the use of AI are transborder, requiring international cooperation (OECD, 2019; UNESCO, 2021).

Finally, the results indicate that values such as fairness, transparency, explainability, accountability, protection of privacy and respect for human dignity could be taken into account regarding the responsible governance of AI. Technological innovations that are developed must always be supported by ethical guidelines and laws which will ensure that these innovations are developed responsibly and the people that are going to use them are not misled. For organisations developing AI, ethical assessment, involvement of various stakeholders and ongoing evaluation should be an integral part of the design, implementation and deployment of intelligent technologies. These strategies are essential in order to build sustainable, inclusive, and equitable societal development in the presence of AI rather than deepening inequalities or creating new ones.

While the study provides a comprehensive interdisciplinary synthesis of the knowledge, it is important to note that there are some limitations. Firstly, the research conducted is only qualitative and no primary data (e.g. interview, focus group and participant observation) are included in this study. Document analysis is suitable for cross-disciplinary and cross-institutional synthesis, but firsthand experience from educators, policy makers, techies and citizens might yield deeper insights of lived experiences of AI uptake. Third, while the international literature does exist, it is important to recognize that much of the current research on AI scholarship is from developed countries, where institutional capacities, digital infrastructure and governance structures are vastly different from those in developing regions, raising questions about the extent to which the results of this research will

apply to the potential implementation of AI in these areas. Third, the field of AI continues to evolve, as do new innovations, resulting in knowledge, public policies and technological capabilities that need to be regularly revisited and re-evaluated.

This study has some important limitations, which present possibilities for future qualitative research. Qualitative strategies such as multiple case studies, interviews, or focus groups could be employed to deepen understanding of the experiences of different stakeholders in the context of the different transformations of the educational institution, public institutions, healthcare and workplace institutions and their experiences of AI. Comparative qualitative studies worldwide, across countries and between developed and developing countries, can help build the understanding of AI governance, digital inclusion and institutional adaptation that is more globally representative. Researchers of the future can also investigate new challenges that come with the use of generative AI, algorithmic governance, digital citizenship, cybersecurity, environmental sustainability, and the long-term psychological and social impacts of greater interaction humans with intelligent technologies. Interdisciplinary research collaborations between social scientists, educators, policymakers, legal scholars, ethicists and computer scientists are necessary to underpin more comprehensive theory that can address the evolving and dynamic relationship between AI and society.

To sum up, the qualitative study we conducted reveals that Artificial Intelligence and digital transformation is not just about technology, but it is about a social process that is transforming the learning, communication, working, governing and engaging with the modern world. The applications of AI in education, public sector, organisational productivity, scientific research, and knowledge dissemination are vast. However, these opportunities also have their own set of issues, including ethical, governance, privacy, digital inequality, transparency, trust and social justice issues. The interdisciplinary,

human-centred and ethical approach to technology must be taken to address these challenges, treating technology as a tool for human well-being, rather than an objective.

Last but not least, decisions made by governments, educators, institutions, researchers and society in general will play a part in defining the role of AI in the future, and not just because of the complexity of the algorithms or the speed of technology advancement. However, responsible governance, inclusive public policies, ethical technological development, and constant questioning and critique will shape how AI can drive equitable and sustainable social progress. This study, thus, with an interpretivist

approach (qualitative), emphasizes the social, institutional and human aspects of AI, which contributes to the existing interdisciplinary studies on digital transformation. In this interpretivist approach (qualitative), the study emphasizes the social, institutional and human aspects of AI, and it is a contribution to the growing interdisciplinary literature regarding digital transformation. Lastly, the adoption of innovative solutions rooted in transparency, accountability, respect for human dignity, and shared objectives of improving the quality of life for current and future generations is essential in supporting the responsible integration of AI into society.

References

- Acemoglu, D., & Restrepo, P. (2020). Robots and jobs: Evidence from US labor markets. *Journal of Political Economy*, 128(6), 2188–2244. <https://doi.org/10.1086/705716>
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.
- Coeckelbergh, M. (2020). *AI ethics*. MIT Press.
- Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). Sage Publications.
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M., Al-Busaidi, K. A., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhalis, D., ... Wright, R. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Floridi, L., & Cows, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Holmes, W., Bialik, M., & Fadel, C. (2022). *Artificial intelligence in education: Promises and implications for teaching and learning* (Updated ed.). Center for Curriculum Redesign.
- Lincoln, Y. S., Guba, E. G., & Lincoln, Y. S. (1985). *Naturalistic inquiry*. Sage Publications.
- OECD. (2019). *OECD principles on artificial intelligence*. <https://oecd.ai/en/ai-principles>
- Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). Sage Publications.
- Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4th ed.). Pearson.
- Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). Sage Publications.
- UNESCO. (2021). *Recommendation on the ethics of artificial intelligence*. UNESCO. <https://unesdoc.unesco.org/>

-
- UNESCO. (2023). *Global education monitoring report 2023: Technology in education—A tool on whose terms?* UNESCO. <https://www.unesco.org/gem-report>
- van Dijk, J. A. G. M. (2020). *The digital divide*. Polity Press.
- Vial, G. (2019). Understanding digital transformation: A review and a research agenda. *The Journal of Strategic Information Systems*, 28(2), 118–144. <https://doi.org/10.1016/j.jsis.2019.01.003>
- World Bank. (2021). *World development report 2021: Data for better lives*. World Bank. <https://www.worldbank.org/en/publication/wdr2021>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.