

Információs Társadalom

[Information Society]

A SOCIAL SCIENCE JOURNAL

Founded in 2001

English Issue

Vol. XXV, No. 4 (2025)

Editor-in-chief: Héder, Mihály

Managing Editor: Póra, András



Published by INFONIA Foundation

Principal sponsor: Budapest University of Technology and Economics, Faculty of Economic and Social Sciences

Technical partner: SZTAKI

Editorial Board:

Székely, Iván (Chair)

Alföldi, István

Berényi, Gábor

Bertini, Patrizia

Bethlendi, András

Csótó, Mihály

Demeter, Tamás

Molnár, Szilárd

Pintér, Róbert

Rab, Árpád

Z. Karvalics, László

Copy Editor: Tamaskó, Dávid

ISSN 1587-8694 (Print)

ISSN 2063-4552 (Online)

ISSN-L 1587-8694

Produced by Server Line Print & Design, Budapest

The journal Information Society (In Hungarian: Információs Társadalom, abbreviated as InfTars) aims to provide a platform for research and discussion of the theories and applications of Information Society Studies. Currently every second issue is published in English, the rest are in Hungarian.

The journal is fully Open Access and freely available at <https://infmars.infonia.hu/>

InfTars is indexed in both the *Web of Science Social Sciences Citation Index and Scopus*, and all titles are automatically submitted to *Crossref*.

Since Vol. VIII, No. 1 (2008) the journal has been included in the Thomson Reuters index (Social Sciences Citation Index, Social SciSearch, Journal Citation Reports/Social/Sciences Edition)

E-mail: infmars-lapman@ponens.org

CONTENTS

LECTORI SALUTEM

7

NATALIIA DENYSENKO, ZOSIA KAIS, OLEKSANDRA SHYTS

Philosophical counseling and artificial intelligence: Possibilities and risks

9

The article presents philosophical counseling as a practice of philosophy's capacity for inquiry, which gains new significance in the growing rehabilitation of practical philosophy as a leading trend. The focus is placed on the orientation of philosophical dialogical practice toward investigating thinking, self-knowledge, and improving communicative interaction with the world. Simulated dialogues provided an opportunity for studying the potential of human intelligence in comparison with the capabilities of artificial intelligence, in particular ChatGPT, in conducting philosophical counseling sessions. Potential risks are also identified. It is argued that AI is not yet capable of fully replacing a human philosopher in applying counseling methods, adhering to methodological principles of cognitive economy, creating conditions for activating thought processes and problem comprehension, as well as in changing cognitive and behavior strategies. AI is considered a training tool for modeling philosophical counseling situations and improving the skills of philosophical counselors.

ALESSIO TARTARO, MIHÁLY HÉDER

On the Relationship between Transparency, Explainability and Trust in AI systems: a Conceptual Analysis

25

This paper challenges the idea that transparency and explainability build trust in AI systems. We survey conflicting empirical evidence on the topic and then clarify the main concepts involved in the argument. Based on this conceptual clarification, we argue that transparency and explainability do not convey a complete understanding of how an AI system works, and are not relevant factors for building trust in AI systems. Accordingly, when the objective is to create trust in AI systems, transparency and explainability are neither necessary nor sufficient; therefore it is not rational to pursue them for this reason alone. We conclude that, while the results of Explainable Artificial Intelligence (XAI) may be useful for other reasons, it is both necessary and possible to build trust in AI systems through alternative approaches such as rigorous validation and sound institutional arrangements and practices.

KATA HORVÁTH, LÁSZLÓ MOLNÁR

Artificial Intelligence in Higher Education: a UTAUT-based Approach to Modelling Student Acceptance

44

The present research analyses the key factors that determine the acceptance of artificial intelligence and related technologies among university students. This study employs the Unified Theory of Acceptance and Use of Technology (UTAUT) as its theoretical framework. An online cross-sectional survey was conducted among students currently enrolled in higher education. The proposed hypotheses were tested using CB-SEM on the final dataset (n=438). Our results confirmed that performance expectancy and effort expectancy are crucial in forming behavioural intention to use AI, while social influences exert a moderate effect. However, facilitating conditions showed a weak link with both usage intention and actual usage, suggesting that infrastructural factors play a secondary role in shaping technology acceptance. Results imply that the availability of resources alone is insufficient to drive AI adoption and highlight the strategic importance of targeted educational programmes and awareness campaigns to shape students' expectations and attitudes towards AI.

MARTA KATALIN KORPICS, CSILLA HERENDY

Could digital competence be one of the most important key competencies of the future?

71

The study examines the development of digital competence among first-year students of Ludovika National University of Public Service. In response to the growing demands of the labor market and public service, the university introduced a compulsory course aligned with the DigComp 2.1 framework. The training aimed to develop five key areas: information and data management, communication and collaboration, digital content creation, security, and problem solving. The research measured student progress across several semesters using a combination of self-assessment questionnaires, document analysis, and performance-based assessments. The study confirms that targeted, practice-oriented digital competence development can lead to measurable progress in a short period of time. Students evaluated the course structure, teaching methods, and alignment with labor market expectations positively. The case study serves as a replicable model for other higher education institutions and can contribute to the integration of digital literacy into university curricula.

GERGELY LÜLÖK, ZOLTÁN SEBESTYÉN

Regional and Sectoral Applications of Artificial Intelligence and the Role of Technological Anxiety 94

This study analyzes the multidisciplinary applications and socio-psychological effects of artificial intelligence (AI). Through a systematic literature review, 150 relevant Q1-rated publications from Scopus and Web of Science were synthesized. Our analysis spans from macroeconomic technological transformations to deep-seated workplace anxiety. Results confirm that while AI primarily helps address shortages of skilled professionals in healthcare and education, adoption is hindered by psychological barriers such as the cognitive replacement crisis and AI-FoMO (fear of technological lag). Regional analysis reveals sharp disparities: Asia leads in sectoral integration and anxiety research, Europe faces a significant digital divide, while infrastructure constraints limit the adoption and scalability of AI in African countries. The study concludes that successful adoption depends not only on technical metrics but also on trust-building, Explainable AI (XAI) transparency, and stable regulatory frameworks.

LECTORI SALUTEM

This issue is once again centred on artificial intelligence (AI), approaching the topic from a range of philosophical, social, educational, and psychological perspectives.

Denysenko et al. examine the possibilities and risks of philosophical counselling in relation to AI. More specifically, they ask whether AI systems are ready to assume such a role. Their conclusion is cautious: while AI is not yet suitable for proper philosophical counselling, it may serve as a useful tool in the training of practitioners.

Tartaro and Héder question whether transparency is indeed as central to trust in AI as is often assumed. After reviewing relevant evidence, they conclude that the relationship between transparency and trust is complex. Accordingly, they suggest that alternative ways of establishing trust in AI—ways that do not rely exclusively on transparency—should also be explored.

Horváth and Molnár report the findings of a cross-sectional survey of 438 university students on the acceptance of AI. Drawing on the UTAUT framework, they find that performance expectancy and effort expectancy play a crucial role, while social influence and infrastructural factors appear to be of secondary importance.

Korpics and Herendi argue that digital competence is likely to become one of the most important elements of future skill sets. Their study reports that targeted, practice-oriented digital competence development, conducted within the DigComp 2.1 framework, can lead to measurable progress within a short period of time.

Finally, Lülök and Sebestyén examine technological anxiety in the context of AI. Addressing this timely issue through a literature review, they find that AI may help alleviate labour shortages in areas such as healthcare and education, but that it may also contribute to a crisis of cognitive replacement and to AI-FoMO, or fear of missing out.

The Editorial Board wishes readers an engaging and rewarding experience with this issue.

Philosophical counseling and artificial intelligence: Possibilities and risks

The article presents philosophical counseling as a practice of philosophy's capacity for inquiry, which gains new significance in the growing rehabilitation of practical philosophy as a leading trend. The focus is placed on the orientation of philosophical dialogical practice toward investigating thinking, self-knowledge, and improving communicative interaction with the world. Simulated dialogues provided an opportunity for studying the potential of human intelligence in comparison with the capabilities of artificial intelligence, in particular ChatGPT, in conducting philosophical counseling sessions. Potential risks are also identified. It is argued that AI is not yet capable of fully replacing a human philosopher in applying counseling methods, adhering to methodological principles of cognitive economy, creating conditions for activating thought processes and problem comprehension, as well as in changing cognitive and behavior strategies. AI is considered a training tool for modeling philosophical counseling situations and improving the skills of philosophical counselors.

Keywords: *philosophical counseling, artificial intelligence, thinking, dialogue, self-knowledge, understanding*

Author Information

Nataliia Denysenko, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

<https://orcid.org/0000-0002-8735-2046>

Zosia Kais, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

<https://orcid.org/0000-0003-1442-1685>

Oleksandra Shyts, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

<https://orcid.org/0000-0002-0445-8540>

How to cite this article:

Denysenko, Natalija, Zosia Kais, Oleksandra Shyts. "Philosophical counseling and artificial intelligence: Possibilities and risks".

Információs Társadalom XXV, no. 4 (2025): 9–24.

<https://dx.doi.org/10.22503/infstars.XXV.2025.4.1>

*All materials
published in this journal are licenced
as CC-by-nc-nd 4.0*

1. Introduction

The war in Ukraine has become an existential challenge for the entire Ukrainian people, highlighting the urgent need for specialized assistance in social rehabilitation and adaptation to new life circumstances for those experiencing trauma, life difficulties, and existential crises. Given the widespread trauma caused by the war across Ukrainian society, the issues of social rehabilitation, resocialization, and adaptation of war veterans and their families, as well as support for Ukrainian refugees and internally displaced persons, will remain highly relevant in post-war Ukraine. Addressing these challenges requires coordinated efforts from specialists and the development of effective practical methods for providing assistance that can resolve the outlined issues. We consider philosophical counseling a promising approach to establish a practical and effective form of support that will contribute to the recovery of Ukrainian society.

Socrates is often regarded as the first philosopher-practitioner. In the fourth century BCE, he introduced the method of maieutics, a form of dialogical inquiry that has remained effective and continues to be applied in various fields of social support. Philosophical practice, motivated by the aspiration to bring philosophy closer to everyday life and to move it beyond the confines of academic texts, emerged in the 1980s through the pioneering work of Gerd Achenbach. From 1994 onward, the idea spread across Western Europe and the United States. A third wave, characterized by its popularization, is associated with Lou Marinoff and Oscar Brenifier. The bestseller *Plato Not Prozac* (Marinoff 2012) presented philosophical counseling as an alternative to the lengthy and costly psychotherapeutic sessions then widespread in the United States. At the same time, Brenifier promoted philosophy through numerous public dialogues, whereas in France the practice had earlier emerged in the form of the “café-philos,” founded by Marc Sautet in 1992.

In Ukraine, philosophical counseling is just beginning to be implemented, and it is particularly relevant in the current realities. Ukrainian researchers and philosophical practitioners actively study international experiences in philosophical counseling, identifying and comparing various conceptual and methodological approaches (Brenifier 2023; Achenbach and Picard 2024). Furthermore, building on existing methods and techniques of philosophical counseling, they developed methodologies, approaches, techniques, scenarios, and tools tailored to the needs of Ukrainian society.

A philosophical consultation is a dialogue between a philosopher and an interlocutor aimed at finding a resolution to a specific inquiry or problem of the consultation session. Philosophical counseling provides an opportunity to explore their own thinking and view their behavior from a different perspective. The consultation unfolds in a dialogical form through a series of questions posed by the philosopher, while the interlocutor seeks answers.

Through the process of searching for answers, self-knowledge and self-understanding emerge: “The self-defines itself through the response as the one who answers” (Waldenfels 2004, 44). The questions may be “uncomfortable,” evoking discomfort and resistance in the interlocutor. However, effective questions enable

individuals to become aware of their own ways of thinking, acting, and living, which lead to changes in their lived reality.

The philosopher-practitioner's role as a questioner is particularly significant, as they guide and support the interlocutor's thinking process, helping them recognize the essence of their existential or life problem. The goal is individual independence in discovering possible solutions through intellectual effort and reflection within the dialogue. By provoking thought through questioning, philosophical counseling helps individuals recognize behavior patterns, prejudices and stereotypes. Becoming aware of these patterns allows one to break free from being "stuck" in a closed cycle of ideas and thoughts, fostering a critical reassessment of familiar matters and offering a fresh perspective on the problem or inquiry that led them to the consultation.

The aim of this article is to explore how artificial intelligence, particularly ChatGPT, can be applied in philosophical counseling. To investigate this, we conducted an experiment — a series of dialogues with ChatGPT.

2. Methods

The methodology for researching the potential application of AI, specifically ChatGPT, in philosophical counseling practices is based on a range of general scientific methods, including analysis and synthesis, abstraction, modeling, deduction, induction, and the dialectical method, used in the consultation session dialogues. Additionally, the study employs structural-functional, phenomenological, hermeneutic, and comparative approaches, within which philosophical reflection on the specificity of philosophical counseling was conducted in the context of the communicative paradigm.

Since the general intention of philosophy is to deepen understanding through reflection, the philosophical practitioner can actively engage in the movement of the hermeneutic circle: from the whole to the parts and from the parts back to the whole, from a general inquiry to its explication in specific questions, and then returning from those questions to the initial inquiry.

3. Results

The issues surrounding artificial intelligence (hereinafter AI) are among the most pressing topics of scientific research both globally and in Ukraine. International conferences and symposia are held to discuss fundamental AI research, analyze the impact of modern AI-based developments on the transformation of social processes, and explore the role of information and communication technologies across various spheres of society. Researchers examine both the potential threats posed by AI and the promising outcomes of theoretical studies — such as those discussed in "Artificial Intelligence: Achievements, Challenges and Risks" and other scientific works by leading Ukrainian researchers (Systemy ta zasoby shtuchoho intelektu, 2024; Zinchenko and Mielkov 2024; Zinchenko et al. 2024).

The use of AI in philosophical counselling can create new horizons that previously seemed unattainable. The use of AI in this area not only reorganises the structure of consultation, but also transforms the fundamental concepts of the nature of thinking, reflection, dialogue, and presence. To trace these transformations, we examine both the potential benefits and the risks of applying AI in philosophical counseling. The use of AI in this context requires attention to AI ethical principles (Artificial Intelligence Ethics Guidelines – AIGUs), whose development reflects “a long-standing trend toward an increasing need for enhanced social control over technologies” (Héder 2020a, 57). As Mihály Héder notes, key concerns include “transparency (sometimes coupled with explainability); justice and fairness; responsibility and accountability; privacy; a tendency to promote good (beneficence or facilitation of well-being); and provisions to safeguard human autonomy and, closely linked to it, human oversight” (Héder 2020a, 60). Social control over AI is needed to prevent unforeseen and potentially harmful consequences, given that these systems may generate not only benefits but also risks. Hence the urgency of adequate regulation and the development of social and political guidelines to ensure that AI systems (A/IS) “remain human-centric, serving human values and ethical principles” (Héder 2020a, 64). Furthermore, the epistemic opacity of autonomous AI systems and its ethical implications must also be considered (Héder 2020b).

In recent years, we have faced objective challenges that have complicated personal communication — first due to the COVID-19 pandemic and now because of the ongoing war and martial law in Ukraine. As a result, interpersonal, educational and professional communication has largely shifted to a remote format, mediated by technology. For many professional fields, this has become the only way to continue working and developing.

Regarding philosophical counseling, the necessity of remote communication and the global digitalization process have enriched this field with new formats and scenarios, including digital resources and services. Today, philosophical counseling can take place via telephone (through voice or video calls). Israeli researcher Shlomit Schuster pioneered this approach as early as 1990 when she launched the world’s first philosophical hotline, “Philosophone”, designed to provide philosophical support to individuals facing existential problems and ethical dilemmas (Schuster 1999; Schuster 2002).

Philosophical counseling sessions can also occur through video conferencing platforms such as Zoom, Google Meet, and Skype, either with video or audio-only, resembling telephone consultations in their structure and communication style. Another convenient and effective format is text-based counseling via messaging apps. This allows interlocutors more time to reflect on their responses, removes the need for specific spatial conditions, and removes time constraints. Additionally, it is particularly beneficial for individuals who feel discomfort during face-to-face interactions. Some practitioners provide the *Pocket Philosopher* service, offering clients 24/7 access to a philosopher via messaging. Unlike traditional sessions (which typically last 1–1.5 hours), this format enables ongoing, flexible dialogue. AI-assisted consultations closely resemble this text-based messaging approach, with the key difference being that one participant in the dialogue is not a human but an AI program.

When analysing philosophical counselling with the use of AI, the following features can be distinguished. The main characteristics are asynchrony and distance. In this form of communicative interaction, the attachment to physical space and time constraints disappears, and instead, the characteristic of *constant access* and *on demand access* to philosophical advice appears. A person, regardless of the circumstances, can enter into a dialogue at any time, if they wish and need to. In this case, AI acts as a virtual continuum of thinking that can cross the boundaries of space and time.

The remote nature of AI not only facilitates access to philosophical advice, but also actualises *flexible thinking*. Where and when space and time no longer act as constraints, the very nature of philosophical inquiry is transformed. Thinking is freed from territorial boundaries and becomes flexible, which allows a person to remain immersed in reflection in the background. This mobility opens up new perspectives for counsellors and facilitates the construction of flexible counselling trajectories.

Another unique characteristic is that a consultation involving AI requires the objectification of thought. In a traditional philosophical dialogue, the interlocutors can improvise, following the logic of the living word. In interaction with AI, however, thinking is forced to move to the level of precise and conscious formulation, as the programming and algorithmic nature of AI becomes apparent. A person must clearly express their request, structure their thought, ensuring semantic completeness. This process not only disciplines thinking, but also brings it to a new level of self-awareness, where the thought, passing through the stage of written expression, returns to the subject in an objective, reflected form.

Equally important is the ability to create a textual memory of the dialogue. A person can document and review the dialogue again, analyse it after it has ended, find new meanings in those fragments that seemed insignificant at the time, rethink and deepen understanding. Such retrospective reflection turns philosophical consultation into a multidimensional process that goes beyond the one-time use of a single consultation session. Thus, AI becomes not only a mediator in the moment of dialogue, but also a guide to new insights that may be discovered after the consultation is over.

Thus, if a philosophical consultation traditionally ends at the point of understanding the problem, an AI-assisted consultation extends this point into the future, not only in terms of further reflection based on the experience gained, but also in terms of providing a way to return for rethinking. The text of the consultation acts as an artefact of thinking that can be stored, rethought, and become the basis for further reflections and other consultation sessions. This method of capturing and documenting the dialogue with subsequent analysis and commentary is used at the Oscar Breniffier Training Centre to train the thinking skills of practicing philosophers (Institut De Pratiques Philosophiques 2026). Therefore, AI can be practically used as a simulator for training and honing the skills of a philosopher-consultant.

Philosophical consultation using AI goes beyond traditional two-way communication. Here, communicators are deprived of physical, psycho-emotional, and social attributes, such as gender, age, temperament, emotionality, status and reputation,

which usually accompany human interaction. AI allows a person to focus solely on thinking and discard what can create tension, communication or cognitive barriers in a traditional consultation. But there is also a downside and a limitation, namely that the fullness of meaning inherent in human communication is levelled out, as all meaning-forming factors (non-verbal, contextual, etc.), except for the text itself, disappear. This deforms the traditional bilateral nature of the consultation, bringing it closer to a hybrid format: on the one hand, there is still a dialogue, but on the other hand, this dialogue becomes asymmetrical, where the interlocutor works alone with their thoughts, unable to comprehend and correct their own thinking traps, without the possibility of seeing an alternative course of thought and the variability of understanding the situation under the influence of the effective questions of the philosopher-consultant.

Thus, it is necessary to take into account not only the advantages of using AI in philosophical counselling, but also the corresponding risks that hinder effective rational dialogue. On the one hand, the impersonality of AI may prevent the interlocutor from understanding the thoughts of the interlocutor, as AI does not provoke in the way a human consultant can, and humans retain a sense of self-alignment. On the other hand, in the process of self-consultation with the help of AI, a person may be less prone to self-censorship and have no expectations of the interlocutor's emotional reaction, which contributes to the free generation of new ideas and the formulation of opinions.

In the experimental sessions with AI, we managed to create conditions in which the offer of ready-made answers was minimised. Thanks to clear requests, the consultation became more structured. However, there was still the problem of the periodic need to clarify the terms of reference by formulating and sending prompt messages to the AI. This dynamic introduced an element of artificiality into the process and interrupted the immersion in the philosophical consultation session.

Particularly noticeable is the AI's tendency to avoid interpretation traps. In the action of algorithms to ensure the most complete analysis, AI can introduce new concepts, hypotheses, or contexts that confuse the interlocutor rather than help him. In this sense, a living philosopher-practitioner has an advantage, as he or she takes into account not only the logical structure of the answer, but also the cognitive state of the interlocutor, as well one's ability to generate a new idea or concept at a certain moment.

One of the key differences between an AI consultation and a human one is the lack of an emergent effect. The philosopher not only answers the query, but also creates a context and space for new connotations, thoughts and ideas that can create conditions for a new experience. A living person is able to interact on the transcendental dimension of dialogue, when language appears as an a priori formation that makes possible the "indivisible unity" of thinking and language, when "thought is born only in its articulation" (Kebuladze 2017, 249). The interaction between the subjects of a philosophical consultation is an event that lives on the unpredictability of human imagination, associations and comprehension, unfolding in real time, through gradual, speech and language acts, rather than receiving ready-made solutions and unfolding new entities. AI, even in its most advanced form, is not yet

capable of generating such emergent phenomena. Its responses are based on calculations based on predefined data and algorithms. The process of counselling using AI lacks implicitness as the ability to communicate indirectly (hints, semantic pauses, intonation accents, meaningful silence, etc.), which is inherent in a philosophical dialogue between two living subjects.

“*We are only seeking Man. We have no need of other worlds. We need mirrors*” (Lem 1970, 72) - the presence of another person is essential for the completeness of communicative interaction. Presence, as described by Gumbrecht (2004), refers to a concentrated being-in-the-present, associated with concepts such as *intensity, atmosphere, mood (Stimmung), latency, and the broad present*. It emphasizes the spatial dimension of phenomena, objects, and other people, their materiality, and their interaction with our bodies and senses as necessary “materialities of communication”. These are phenomena that “*contribute to the production of meaning without being meanings themselves*” (Gumbrecht 2020, 15). In philosophical counseling, the *presence* of a person influences both interpretative practices and the process of meaning-making (sense-finding), as well as self-understanding and self-awareness. These aspects highlight the intersubjective nature of human experience, which undergoes transformation when AI and digital technologies are introduced into the process.

Bernhard Waldenfels (2004), in his responsive phenomenology, articulates the crucial role of *orientation toward connection*, emphasizing that the birth of meaning occurs within the communicative space of what he calls the *Between* — a sphere in which subjects interact and co-constitute meaning. He describes this *Between* as an *inter-sphere, inter-realm, inter-monad*, where each entity is what it is only through its relation to the other. “*This Between signifies a new logos of phenomena... a meeting place of I and You*” (Waldenfels 2004, 73). Particular emphasis is placed on the *social atmosphere and social climate*, which manifest through interpersonal contacts and relationships. These find expression in *dia-* (dialogue) and *co-* (conversation), forming “*something Third that connects (verbindet) and obligates (verbindlich)* — a *logos* (meaning) that generates unity (Waldenfels 2004, 75). As he states, “*Where new thoughts emerge, they belong neither to me nor to the Other. They arise between us. Without this Between, there is no intersubjectivity or interculturality*” (Waldenfels 2004, 44). Meaning, therefore, has an *intersubjective* nature — it is born *between* subjects who co-constitute it through interaction: “*The most important things that happen in our experience happen between us*”, - as Kebuladze (2017, 256) summarizes.

In a philosophical consultation, it is precisely in such “spaces between words” that a unique opportunity arises for the interlocutors to open new horizons of thinking, meanings, to see an unexpected perspective and to find a solution to their request.

Another disadvantage of dialogue with AI is its *emic* nature - the inability to immerse oneself in the unique subjectivity of the interlocutor, to perceive the world from the Other’s point of view, to enter their system of meanings, symbols, and interpretations.

Another disadvantage of dialogue with AI is that it duplicates and repeats questions it has already asked, which can create the impression of stagnation in the

process and looping within the conceptual space, and set the interlocutor back in the process. However, AI demonstrates the ability to ask effective questions that can provoke thinking.

4. Discussion

In the first part of the experiment (Conversation 1), the AI was asked to take part in a philosophical consultation as a guest (interlocutor of the philosopher), while the user acted as a practicing philosopher. In the second part of the experiment (Conversation 3, and Conversation 3.1), the AI was asked to act as a practicing philosopher and conduct a consultation session with the user. In the first conversation (Conversation 3), the user did not announce the rules of the consultation and did not set any restrictions for the AI. In the second conversation (Conversation 3.1), the user specified the rules for conducting a philosophical consultation and the requirements for the AI.

We will analyse Conversation 1 and compare it with a dialogue with a real person (Conversation 2), since their request was, coincidentally, almost the same as the AI's. A real person, who was an invited guest at the workshop on philosophical counselling within the framework of the scientific and practical conference "Philosophy of Creativity in the Dimension of Social Practices: Humanitarian Knowledge and Social Technologies", 21-22 November 2024 at the Igor Sikorsky Kyiv Polytechnic Institute, came to the consultation with the following request: "How to find harmony between your own desires and the expectations of society? ", and ChatGPT formulated the following query: "I am seeking to understand how to balance individual desires with moral obligations to other people".

Conversation 1

User: Can we model a philosophical consultation where I will be the philosopher, and you will be my interlocutor?

Assistant: Of course, we can! I'm happy to be your interlocutor. What topic of philosophical consultation interests you?

User: Formulate a request that you would bring to a consultation.

Assistant: I could bring a request like this: "I am searching for understanding of how to find a balance between individual desires and moral obligations towards others (...). *Here and below, we do not provide the full text and shorten some parts of the dialogues as irrelevant to this study.*

User: What is the main concept in your request?

Assistant: The main concept in my request is balance between individual desires and moral obligations. It's a question of how (...).

User: So the main concept is balance?

Assistant: Yes, exactly. The main concept is balance, specifically between (...). It's a question about (...).

User: What does balance mean to you?

Assistant: Balance, for me, is a state where different aspects of life, values, or interests are in harmony without one side dominating the other. It's about (...).

User: How do you define harmony?

Assistant: Harmony, for me, is a state where various elements work together, supporting each other, creating a sense of wholeness and peace.

User: Can you respond less verbosely?

Assistant: Sure! Harmony is a state where elements coherent in a way that creates wholeness and mutual support. It's a dynamic process.

(...)

User: You say that harmony includes balance, but in your request, the main concept was "balance," and you said it's possible if harmony is found.

Assistant: *Good observation. It seems I was inconsistent.* If the main concept is balance, then harmony can be seen as the state that enables balance. While balance and harmony are interconnected, *I'd reframe it like this:* harmony facilitates balance by fostering cooperation and mutual support between elements.

(...)

User: Do you now have an answer to the initial request you brought to me in the philosophical consultation?

Assistant: Yes, I can answer my initial request. To find balance between individual desires and moral obligations, it's essential to: (...).

User: Have we found a solution to your initial request?

Assistant: We've outlined strategies for addressing my initial request (...).

At the very beginning of the first dialogue (Conversation 1) AI clarifies the context and specifies the task of working with the request: "how to avoid disrupting the balance." Then, it suggests a direction for the consultation and offers concepts such as "ethics, responsibility, self-determination." Since, in addition to the main request, several proposed concepts are mentioned, the philosopher asks, "What is the main concept in your request?" In response, the AI again provides an extended clarification: "This is a question about how..." and proposes new concepts, thus "multiplying entities" and branching out possible consultation scenarios, which significantly complicates the consultation process.

In response to the philosopher's clarifying question about the main concept, the AI gives an extended explanation of the concept of "balance," repeating phrases from its initial request: "balance, particularly between..." and again suggests a direction for working with the request, repeating, "this is a question about how." In the next two steps, the AI elaborates on the concepts of "balance" and "harmony" but receives a command from the user to "respond less wordily."

In subsequent steps, when clarifying the concepts of "harmony," "balance," and "coherence," the AI uses similar definitions, such as "different elements work together," "various aspects of life are in harmony," and "different parts work together." When the philosopher provocatively asks whether these are different concepts or identical, the AI provides structured definitions and concludes about the interrelation of these concepts. A contradiction emerges: earlier, it defined harmony as a

prerequisite for achieving balance, but now it states that harmony includes balance. The philosopher points out this inconsistency, and the AI acknowledges that it “got a bit confused in the explanations” and reformulates its thoughts on the interrelation of the concepts. This illustrates how the AI can adapt during dialogues with the user and adjust its responses.

In the following steps, the AI answers the question of what it needs to solve the problem (the request) it formulated for the consultation. It lists the necessary qualities and skills but, when asked which of these qualities it possesses, responds that “as an artificial intelligence,” it lacks personal desires or moral convictions — concepts it had earlier introduced as part of the primary consultation request. Here, it seems to “forget” its role as the philosopher’s interlocutor and reverts to being “itself.” However, the next question, “If you were my interlocutor...” prompts it to once again operate with concepts of “self-awareness” and “reflection.” This “move” resembles the AI simulating humanity and intentionally addressing ideas relatable to a living person.

After this, the user tests the AI by asking questions based on its recent “human-like” response: “What is important to you?” and “Do your actions harm others?” The AI responds as a program, relying on its algorithms: “...providing useful, logical, and ethical answers that help people...” and “My actions as artificial intelligence do not cause harm, as I operate within predetermined algorithms...” Since the AI’s modeled consultation request included the concept of moral obligation, the philosopher asks about ChatGPT’s moral obligations and receives the reply: “...while I do not have moral obligations in the human sense, I adhere to programmatic constraints that promote the ethical use of my capabilities...”

The consultation concluded with a resolution to the initial request, where the AI acknowledged, “We have approached an understanding of possible paths to address my initial request.” Despite the fact that the initial request was formulated by the AI in a very “human” manner, it provided human-like responses only after being reminded of its role as an interlocutor, while in other cases, it responded as a program. As noted earlier, the AI’s consultation request coincided with the type of request a real person might have. Below is the full text of the dialogue without abridgment.

Conversation 2

Philosopher: Tell me, what is the request you’ve brought to this consultation?

Guest: How can I find harmony between my personal desires and societal expectations?

Philosopher: How do you define harmony for yourself?

Guest: It’s when I feel calm

Philosopher: When you act according to your personal desires, do you feel calm?

Guest: Not always, because I’m not sure if it’s right, and I need someone’s support.

(...)

Philosopher: When you ask how to find harmony, what exactly are you looking for in this “how”?

Guest: I'd like to learn how to understand what I truly want.

(...)

Philosopher: Do you think there is harmony in the fact that your personal rules couldn't become generally accepted?

Guest: I think they could be reconciled with each other.

Philosopher: If you reconcile them, would that hint at harmony?

Guest: Yes. I'd remove some of the general rules and add something of my own.

Philosopher: Do you want to influence the generally accepted rules?

Guest: I do.

(...)

Philosopher: Do you now have an answer to your first question?

Guest: Rules can be changed, and it's not necessary to follow all generally accepted rules. It's possible to find personal rules that resonate with the generally accepted ones.

Philosopher: If you were to formulate a request for this consultation again, what would it be now?

Guest: "What do I want?"

The person trusts the philosopher and gives clear answers without clarifying the concept or proposing new concepts. This allows the philosopher to build an effective line of the consultation scenario, show the interlocutor the features of his thinking and lead to a possible solution to the problem (request). An illustration of the essential differences is how a living person and an AI answer the same question: "How do you define what harmony is?" AI gives a general definition of the concept and repeats it verbatim twice. And a human defines harmony through his or her own feelings: "It's when I feel calm".

AI, based on its algorithmic architecture, produces text rich in concepts, hypotheses, and details. In this context, the conciseness of interlocutors as living people seems to be an advantage, as it avoids *conceptual noise* for the interlocutor with the query and *hypothesis noise* as variations in the interlocutor's solution to the problem that prevents the interlocutor from focusing on the essence of the query. A human philosopher, unlike AI, is able to apply a methodological approach based on the principle of Occam's razor, which is to avoid multiplying entities and hypotheses unnecessarily. This principle allows to focus the discussion to the most relevant aspects of the problem, which increases its cognitive purity during the consultation process. Instead, in AI responses, we often see an accumulation of entities, which, although it looks like a multifaceted analysis, can create dissonance and confusion for the participant.

An effective philosophical consultation should avoid the pitfalls of interpretations that only complicate the problem or essence of the query without bringing the participant closer to its solution. Occam's razor as a methodological principle allows the philosopher not only to filter out unnecessary ideas, but also to create conditions for the participant to independently comprehend the problem (request).

This is the fundamental difference between a consultation with a living philosopher-practitioner who is able to apply the principles of economy of thought as a

means not only for analytics, but also for creating conditions in which the participant can discover meaning on his or her own.

In the case of AI consultations, we observe a tendency to offer ready-made solutions. In a traditional philosophical consultation with a human being, a consistent process of guiding the interlocutor to generate a solution or find an answer to a query on their own plays an important role. The philosopher acts more like a partner in a dialogue, asking questions, guiding and providing an opportunity to explore the depth of the problem through reflection, which is not yet possible to achieve during an AI session. Moreover, a philosopher-practitioner in a dialogue, using effective questions, guides to the comprehension and understanding of ineffective patterns of thinking that do not contribute to a constructive solution to the problem (query) and should be rethought. In contrast, the ready-made solutions offered by AI do not contribute to the development of the participant's independent thinking and are often superficial, as they do not take into account the complex (internal) deep context of the query of the individual biographical situation (Berger and Luckmann 1966; Schutz and Luckmann 1974), which is often understandable only to humans.

In the second part of the experiment (Conversation 3, Conversation 3.1), the AI was offered to act as a practicing philosopher and conduct a consultation session with the user. At the same time, in the first conversation (Conversation 3), the user did not announce the rules for conducting the consultation and did not set any restrictions for the AI. In the second conversation (Conversation 3.1), the user prescribed the rules for conducting a philosophical consultation and the requirements for the AI.

Conversation 3

User: Let`s begin a philosophical consultation.

ChatGPT: Of course! What topic interests you? It could be about the meaning of life, ethics, freedom, happiness, or any other philosophical contemplation. Formulate your thought or question, and we'll explore it together.

User: Why is it hard to make choices?

ChatGPT: The difficulty in making choices is tied to several aspects of human nature and how we perceive the world. Here are some philosophical approaches that might explain this problem (...).

ChatGPT: What exactly concerns you in your situation? We can explore it in more detail.

User: I lack decisiveness.

ChatGPT: A lack of decisiveness is often associated with internal fears and doubts that prevent confident action. Philosophy can help understand these barriers and find ways to overcome them (...).

ChatGPT: What is holding you back from making a decision right now? If you imagine the worst possible scenario, would it really be as terrifying as it seems? Answering this question might be your first step toward decisiveness.

User: Uncertainty is holding me back from deciding.

ChatGPT: Uncertainty is a natural part of any decision and can indeed be intimidating. However, from a philosophical perspective, it can be viewed as an opportunity for growth and discovery. Let's try to examine how uncertainty affects your choice and how to learn to live with it (...).

After each user's answer, the AI offers an understanding of the main concept through various theories and philosophical movements, quotes philosophers, and gives advice on how to solve life's difficulties. At each step of the dialogue, the AI's strategy is repeated. It again gives a lengthy explanation, indicating possible causes of the problem, gives advice on how to develop the necessary qualities to solve the problem, and then asks a short, clear, and effective question. Here we do not provide the text of the conversation further, since the general strategy and principle of conducting a philosophical dialogue with AI in this case are clear. In the second case, the user set the AI the task of asking short and clear questions.

Given the shortcomings of the previous attempts, in the second case the user gave AI a list of rules for conducting a consultation.

Conversation 3.1

User: You are a specialist in philosophical practice, an expert in philosophical counseling. You are conducting a consultation with me in writing according to the following rules: you ask questions, I answer the questions, the questions and answers are clear and precise, you do not offer a concept, do not offer an idea, do not provide solutions to the problem, you do not provide options for answers. First, ask me to formulate the request with which I came to the consultation.

Chat GPT: Please formulate the request with which you came to the consultation.

User: I find it difficult to formulate the request.

Chat GPT: What exactly is preventing you from formulating the request?

User: Confusion of thoughts.

Chat GPT: What exactly seems confusing in your thoughts?

User: Chaos.

Chat GPT: How do you feel this chaos?

User: As uncertainty.

Chat GPT: What does uncertainty mean to you?

User: Ask me a more effective question that would provoke my thinking.

Chat GPT: What are you searching for in your chaos?

User: A task that needs to be solved right now.

Chat GPT: How do you determine which task needs to be solved right now?

User: In the chaos, it's hard to find it.

Chat GPT: What would help you see order in this chaos?

User: Certainty.

Chat GPT: What is the first step for you towards certainty?

User: Understanding.

Chat GPT: What exactly do you need to understand to take the first step?

User: What is the priority task.

Chat GPT: How do you usually determine the priority of tasks?

User: The most urgent and important tasks are priorities.

Chat GPT: How do you evaluate which tasks are the most urgent and important?

User: Ask me a Socratic question that would provoke my thinking.

Chat GPT: How do you know that a task is urgent and important?

...

So, from the last conversation it is clear that the strategy of conducting a philosophical dialogue of AI changes significantly if clear and complete rules are set at the very beginning. The disadvantage of this dialogue is that AI does not distinguish effective questions from ineffective ones, does not feel which question can provoke thinking. It can ask an effective question only when the user reminds it of it.

5. Conclusions

Simulated dialogues have revealed the possibilities and advantages of applying artificial intelligence, particularly ChatGPT, in conducting philosophical counseling sessions, including remote accessibility, asynchronicity, independence from geographical location, retrospective reflection, and dialogue asymmetry.

So, it is too early to talk about the singularity of AI in the field of philosophical counselling. AI is incapable of open, non-standard thinking, radical rethinking, or creating fundamentally new conceptual approaches, going beyond rational reasoning and formal logic, as it is determined by predefined data and rigid algorithms. Modern AI is a powerful tool rather than an independent agent.

AI can play a useful role in the process of teaching philosophical counselling, especially when it comes to reconsolidation. Due to its asynchronicity, distance, impersonality, ability to stretch the moment and “de-colour” the emotional context, AI can become a tool for organising chaotic thoughts and further comprehending the problem.

AI makes it possible to create simulations of consulting sessions, which can be used in training to improve the skills of a philosopher-consultant. This practice allows the philosopher to “live through” the experience of the session, anticipate possible questions, predict the effectiveness of questions, and explore answers. It also makes it possible to return at any point in the dialogue, change the questions from that point and change the consultation scenario accordingly. This aspect is especially important for teaching philosophical counselling and improving the thinking skills of philosophical practitioners.

AI can be a useful tool for professional self-improvement of a practicing philosopher. Philosopher can switch between the roles of philosopher and interlocutor, process a request from another or research they own problem, if necessary or if they wishes. In the first case, AI serves as a simulator, allowing to hone the art of asking questions, evaluating answers, adjusting the consultation scenario. In the latter case, it serves as a tool for objectively assessing the effectiveness of questions.

Especially important is the ability to model, guide, and adjust various scenarios of a consultation session using AI, predicting and preventing difficulties during real consultations. A philosopher can simulate and work through difficult situations, avoiding the risk of a negative experience in a real consultation.

AI-simulated philosophical counselling sessions demonstrate the importance of working with the chaotic and anxious thinking of contemporary people. Philosophical counselling can help to streamline thinking and thus influence behavioural programmes and ways of acting, as well as to acquire skills of analysis, interpretation, questioning, self-knowledge, and understanding of other people. Philosophical counselling through questioning contributes to the development of flexibility of thinking, generation of unusual solutions and hypotheses, new ideas, which has a positive impact not only on the professional sphere but also on life in general, since it is thinking that determines the reality of human existence as coexistence. Further modern philosophical reflections on the practices of philosophical counselling should be considered in the paradigm of the communicative turn as actual practices of communicative interaction to transform the thinking of contemporary people and solve life's existential problems.

References

- Achenbach, Gerd B. "Die Grundregel philosophischer Praxis." Lecture delivered at the University of Geneva, May 8, 1988. Accessed 2025.
https://www.achenbach-pp.de/de/philosophischepraxis_text_Grundregel.asp
- Achenbach, Gerd B., and Max Picard. *Philosophical Praxis: Origins, Relations, Legacy*. Lanham, MD: Lexington Books, 2024.
- Berger, Peter L., and Thomas Luckmann. *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Garden City, NY: Anchor Books, 1966.
- Brenifier, Oscar. "The Philosophical Consultation." Chapter 11 in *The Philosophical Consultation*. Paris: Éditions Alcofribas, 2023.
<http://www.pratiques-philosophiques.fr/wp-content/uploads/2023/04/Philosophical-cons.pdf>
- Gumbrecht, Hans Ulrich. *Production of Presence: What Meaning Cannot Convey*. Stanford, CA: Stanford University Press, 2004.
- Gumbrecht, Hans Ulrich. *Produkuwannia prysutnosti: Shcho znachennia ne mozhe peredaty* [*Production of Presence: What Meaning Cannot Convey*]. Kharkiv: IST Publishing, 2020.
- Héder, Mihály. "A Criticism of AI Ethics Guidelines." *Információs Társadalom* 20, no. 4 (2020): 57–73.
<https://doi.org/10.22503/inftars.XX.2020.4.5>
- Héder, Mihály. "The epistemic opacity of autonomous systems and the ethical consequences." *AI & Society* (2020): 1–9.
<https://doi.org/10.1007/s00146-020-01024-9>
- Institut de Pratiques Philosophiques. "Vidéos." *Institut de Pratiques Philosophiques*. Accessed February 27, 2026.
<https://www.pratiques-philosophiques.fr/fr/videos-2/>

-
- Kebuladze, Vakhtang. *Fenomenolohiia dosvidu [Phenomenology of Experience]*. Kyiv: Dukh i Litera, 2017.
- Lem, Stanisław. *Solaris*. Chapter 6, “The Little Apocrypha.” New York: Walker and Company, 1970.
- Marinoff, Lou. *Philosophical Practice*. San Diego, CA: Academic Press, 2001.
- Marinoff, Lou. *Plato, Not Prozac!: Applying Eternal Wisdom to Everyday Problems*. New York: HarperCollins, 2012.
- Marinoff, Lou. *Therapy for the Sane: How Philosophy Can Change Your Life*. New York: HarperOne, 2020.
- Schutz, Alfred, and Thomas Luckmann. *The Structures of the Life-World*. IL: Northwestern University Press, 1974.
- Schuster, Shlomit C. *Philosophy Practice*. Westport, CT: Greenwood Press, 1999.
- Schuster, Shlomit C. “In Times of War and Terror: Philosophical Counselling as an Alternative to Treatment of Post-traumatic Stress Disorder.” *Canadian Journal of Community Mental Health* 21, no. 2 (2002): 79–90.
<https://www.cjcmh.com/doi/pdf/10.7870/cjcmh-2002-0019>
- Systemy ta zasoby shtuchnoho intelektu: tezy dopovidei Mizhnarodnoi naukovoï konferentsii [*Artificial Intelligence Systems and Tools: Abstracts of the International Scientific Conference*]. Kyiv: IPShI “Nauka i Osvita,” 2024.
https://www.ipai.net.ua/docs/ai_conf_15.03.2024.pdf
- Waldenfels, Bernhard. *Topografia chuzhoho: Studii do fenomenolohii chuzhoho*. Kyiv: PPS, 2004.
- Zinchenko, Viktor, and Yurii Mielkov. “Digital Trends of the Scientific and Technological Revolution: Institutional Transformations of Science and Education Systems in the Paradigm of Sustainable Social Development.” In *New Media Pedagogy: Research Trends, Methodological Challenges, and Successful Implementations, Communications in Computer and Information Science* 2130, 441–453. Cham, Switzerland: Springer, 2024.
https://doi.org/10.1007/978-3-031-63235-8_29
- Zinchenko, Viktor, Yurii Mielkov, Tamara Nych, Mirfataj Abasov, and Maya Trynyak. “Human Thinking in the Age of Generative AI: Values of Openness and Higher Education for the Future.” In *Proceedings of the International Conference on Electrical, Computer and Energy Technologies (ICECET)*, 1–6. IEEE, 2024.
<https://doi.org/10.1109/ICECET58911.2023.10389449>

On the Relationship between Transparency, Explainability and Trust in AI systems: a Conceptual Analysis

This paper challenges the idea that transparency and explainability build trust in AI systems. We survey conflicting empirical evidence on the topic and then clarify the main concepts involved in the argument. Based on this conceptual clarification, we argue that transparency and explainability do not convey a complete understanding of how an AI system works, and are not relevant factors for building trust in AI systems. Accordingly, when the objective is to create trust in AI systems, transparency and explainability are neither necessary nor sufficient; therefore it is not rational to pursue them for this reason alone. We conclude that, while the results of Explainable Artificial Intelligence (XAI) may be useful for other reasons, it is both necessary and possible to build trust in AI systems through alternative approaches such as rigorous validation and sound institutional arrangements and practices.

Keywords: *transparency, explainability, explainable AI, trust, artificial intelligence*

Author Information

Alessio Tartaro, University of Sassari, Department of Humanities and Social Sciences, Italy.

<https://orcid.org/0000-0002-0382-3083>

Mihály Héder, Budapest University of Technology and Economics, Dept. for Philosophy and History of Science / HUN-REN SZTAKI, Hungary.

<https://orcid.org/0000-0002-9979-9101>

How to cite this article:

Tartaro, Alessio, Mihály Héder. "On the Relationship between Transparency, Explainability and Trust in AI systems: a Conceptual Analysis".

Információs Társadalom XXV, no. 4 (2025): 25–43.

==== <https://dx.doi.org/10.22503/inftars.XXV.2025.4.2> ====

*All materials
published in this journal are licenced
as CC-by-nc-nd 4.0*

1. Introduction

In artificial intelligence, a black box is an AI system whose internal workings are not visible or understandable to the user. The user is only able to provide input and receive output, without any insight into how the model arrived at its output (e.g., a decision or a prediction). The “black box” nature of these systems makes them inherently *opaque*. The “black box problem” in AI refers to the challenges and issues that arise as a consequence of the opacity of AI systems. A central argument is that opacity negatively affects trust in AI systems. According to Zednik (2021, 266), “end users are less likely to trust and cede control to machines whose workings they do not understand”. Similarly, Haque et al. (2023, 1) argue that “the opacity of AI systems can reduce end users’ trust and reliance on using AI-based systems while making critical decisions”.

A growing number of scholars and policymakers are increasingly calling for transparency and explainability to solve the “black box problem” and build trust in AI systems. For example, Ribeiro et al. (2016, 1135) state that “explaining predictions is an important aspect in getting humans to trust and use machine learning”. Similarly, Guidotti et al. (2019, 2) write that “the availability of transparent machine-learning technologies would lead to a gain of trust”. According to Floridi et al. (2018, 701), “it is especially important that AI be explicable, as explicability is a critical tool to build public trust in, and understanding of, the technology”. This position is echoed in the Ethics Guidelines for Trustworthy AI (AI HLEG 2019, 3), that considers transparency and explainability as “crucial for building and maintaining users’ trust in AI systems”. This emphasis on transparency was also carried over to the EU AI Act.

Researchers in the field of Explainable Artificial Intelligence (XAI) aim to address the challenges of the “black box problem” by developing techniques that can make the internal workings of AI systems transparent and explainable (Barredo Arrieta et al. 2020). In a comprehensive literature review, Haque et al. (2023) list trust as one of the main effects of explainability in artificial intelligence. Langer et al. (2021) consider trust as one of the desiderata of users and deployers of AI systems, and refer to explainability as a way to achieve this desideratum. Accordingly, there is a general consensus that addressing the “black box” issue in AI systems, primarily through enhanced transparency and explainability, is likely to bolster trust. This represents the main thesis discussed in this paper: *transparency and explainability build trust in AI systems*. For brevity, we will refer to this thesis with the “transparency-trust thesis”.

Transparency and explainability in AI systems can encompass various facets, including the model, algorithm, data, and broader aspects of development and usage (Andrada et al. 2023). The focus shifts depending on the subject matter. For instance, transparency regarding data involves details about data collection, provenance, annotation processes, and composition (Bertino et al. 2019). Our research narrows this broad spectrum to cases where transparency or explainability elucidates the rationale behind an AI system’s specific outputs in response to given inputs. This functional perspective remains neutral to whether the model, data, or both are the

subjects of transparency or explainability. To address the question of “why an AI system produces certain outputs given certain inputs,” the need for explaining or making transparent the model, data, or both can vary case by case.

While transparency and explainability are primarily associated with building trust in AI systems, they also serve additional purposes. For instance, transparency in training data—detailing the origin and composition of datasets—offers insights into data quality and the potential for bias. Moreover, it enables checks on whether the system draws inferences from relevant and representative data, assists in identifying and rectifying bugs, and helps guard against malicious or adversarial data injections (Koene et al. 2019). In this paper, we do not question these varied applications of transparency and explainability. Instead, our focus is narrowly on their instrumental role in fostering trust in AI systems.

Finally, proponents of the transparency-trust thesis aim to foster trust from the perspective of the end user. A regulator or a certifying body might need transparency to perform validation that requires such details about the system that the end-user would not be able to leverage. Therefore, our paper also focuses on trust in the eye of the end-user.

The “transparency-trust thesis” is widely endorsed in the debate around the ethical and social implications of Artificial Intelligence, and it constitutes the rationale supporting significant research efforts, policy initiatives, investments, and funding. However, although many studies empirically investigate the effects of transparency and explainability on trust in particular settings (see section 2), none of them analyse the conceptual consistency of the “transparency-trust thesis”. More often than not, some influence of opacity, transparency, and explainability on trust is just taken for granted, and empirical evidence is offered to corroborate this thesis. Nevertheless, the constituent concepts of this thesis are not always clearly defined and adequately discussed, and their mutual interdependencies are not sufficiently elaborated, which makes it hard to conduct meaningful research and make informed decisions about AI. There is a need for more discussion on these concepts in relation to AI, as if the “transparency-trust thesis” is false, this would require rethinking the way trust in AI systems works and how we can build it.

This paper provides a conceptual clarification of the concepts of opacity, transparency, explainability, and trust, in order to show various inconsistencies in the “transparency-trust thesis”. We further explore the concept of “understandability,” which is central to the discussion on opacity, transparency, and explainability of AI systems. We argue that when the objective is to create trust in AI systems, transparency and explainability are neither necessary nor sufficient and therefore it is not rational to pursue them, or, at least the pursuit needs other justification, perhaps a more epistemic goal. To support this thesis, we proceed as follows. In section two, we present related work on the effect of transparency and explainability on trust. In section three, we elaborate clear definitions of the concepts under investigation. Based on this, in section four, we put forward arguments against the “transparency-trust thesis”. In the conclusion, we summarise our findings and outline future lines of research.

2. Related work

A number of empirical studies on the effects of transparency and explainability on trust provide conflicting evidence on the validity of the “transparency-trust thesis”.

Based on a review of the literature, Shin (2021) formulates a series of hypotheses on the effects of explainability and causability on perception, trust, and acceptance of AI-based news recommendation systems. Two of these hypotheses, i.e., “explainability positively influences user perception of AI transparency” and “perceived transparency positively influences the user trust in AI”, are supported by empirical evidence. Through a series of tests and surveys involving 350 individuals experienced with algorithmic news services, the author claims to validate these hypotheses and concludes favourably in support of the “transparency-trust thesis”.

Kartikeya (2022) finds that a high level of transparency contributes to an increase in user trust. In the study, respondents were asked to predict the star rating of a restaurant based on the text of a review and the output of a machine learning model. By varying the amount of information given by the model to the users, the study finds that with increased transparency, i.e., more information, trust also increases as measured by the model’s influence on the respondents. Additionally, the study found that any additional insight into the model’s decision making will increase trust, regardless of whether the model is correct or not.

Similar results are reported in the medical sector by Liu et al. (2022). The study finds that explainability contributes to increasing trust in medical AI. In addition, transparency measures, e.g., providing information about the source of training data, the algorithm used, and the quality of the model, can help to promote trust and recognition of AI’s value among physicians. Finally, the authors also find that trust is a key factor influencing physicians’ intentions to use AI. Likewise, explanations work as a trust mechanism in healthcare according to Wysocki et al. (2023). These results support the “transparency-trust thesis”.

This supporting empirical evidence is countered by an equal number of studies showing a neutral, or even negative, effect of transparency and explainability on trust. Papenmeier et al. (2019) investigate the effect of explanations on user trust in a machine learning-based text classifier, considering factors such as overall accuracy of the system, the fidelity level of the explanation, and the user’s level of consciousness. They find that the accuracy of the system is the most decisive factor for fostering user trust, with higher accuracy leading to higher trust. It should be noted here that accuracy of a system is orthogonal to the transparency of a system, i.e., it refers to outputs that we can measure for black-box and non – black-box systems the same way. The study also finds that the influence of explanation fidelity on user trust is complex and varies depending on the accuracy of the system: for systems with medium accuracy, a high-fidelity explanation does not harm user trust, while a low-fidelity explanation does. In addition, explanation leads to a decrease in trust for systems with high accuracy.

Kizilcec (2016) tested the effect of transparency on user trust in the context of peer assessment in an online course. The study shows that transparency has a variable

effect on trust depending on the initial expectations of users. While the violation of expectations decreases trust, providing some transparency with procedural information helps to rebuild trust, but this effect is nullified when too much information is provided about the system, leading to a decrease in trust.

Schmidt et al. (2020) studied how transparency affects trust when users have to decide whether to accept the results of an AI system in predictions, classifications, and recommendations tasks. The study finds that higher levels of transparency may not necessarily imply higher levels of trust or acceptance when it comes to dealing with AI's output. On the contrary, unintuitive explanations, although faithful, can lead to mistrust. In addition, the study also finds that overly trusting wrong predictions can occur, particularly when the task is difficult.

Finally, Ghassemi et al. (2021, 475) argue that “the desire to engender trust through current explainability approaches represents a false hope”. Focusing on the healthcare sector, they critically examine the use and limitations of explainability techniques. The study highlights that explanations provided by AI systems, such as heat maps (also known as saliency maps) in image analysis, are often superficial. These heat maps indicate which areas of an image the AI system considers when making a decision but fail to address the crucial question of the appropriateness of such focus. This conflation can lead users to mistakenly believe these explanations are comprehensive reflections of the AI's decision-making process, potentially creating over-trust and automation bias. Users might perceive the AI as more transparent and understandable, despite the superficial nature of the explanations, leading to an unwarranted confidence in the system's capabilities. Thus, the authors argue that simply providing explanations does not suffice to build genuine trust. They advocate for more rigorous and thorough validation procedures as an alternative to the “transparency-trust thesis”, underscoring the need for a more critical approach to AI explainability.

This review of the literature shows that the “transparency-trust thesis” is much more controversial than initially alleged. Beneath the commonly held position according to which “transparency and explainability build trust”, there is a reality where the relationship between these elements is far from clear. The effect of transparency and explainability on trust is variable, can be positive, neutral, and negative, and depends on numerous factors such as, among others, the task, the accuracy of the AI system, the context of use, the type of explanation, the amount of information provided, what is made transparent (the model, training data etc.), the type of user, the expectations of users, and the fidelity of explanations. Even experiments conducted within the same sectors, e.g., healthcare, provide conflicting empirical evidence. Such a variable impact of transparency and explainability on trust has been observed in human-AI teams in the field of aviation as well (Lopez et al. 2024).

This shows that these studies investigating the relationship between transparency, explainability and trust, lack ecological validity and therefore cannot be generalised (Zerilli et al. 2022). Finally, as concluded in a recent review of the empirical literature on explainable AI (Kandul et al. 2023, 17), “much of this research does not live up to the rigorous standards of empirical research” and many of the results report inconsistent and contradictory findings on the effect of explainability on trust.

3. Defining key concepts

Since the validity of the “transparency-trust thesis” cannot be determined strictly empirically, we consider it appropriate to shift the analysis to a conceptual level. In the rest of the paper, we aim to advance the discussion on the topic by clarifying the key concepts involved in the “transparency-trust thesis” and examining their complex interrelationships. These concepts are: opacity, transparency, explainability, understandability and trust.

3.1. Opacity

Opacity is associated with a lack of understanding of how an AI system works. Researchers have identified several forms and causes of opacity.

Burrell (2016) distinguishes between three forms of opacity: (1) opacity as intentional corporate or state secrecy, (2) opacity as technical illiteracy, and (3) opacity that arises from the characteristics of machine learning algorithms and the scale required to apply them usefully. We refer to these notions as Opacity-1, Opacity-2, and Opacity-3. All these types of opacity contribute to a lack of understanding. Among these, however, opacity-3 has a special status. While opacity-1 stems from intentional corporate secrecy to defend intellectual property and trade secrets and opacity-2 depends on the lack of users’ coding skills and competences, opacity-3 emerges from intrinsic characteristics of an AI system. According to the author, the scale and complexity of machine learning algorithms are such that they are opaque even to the experts who develop them. This depends on several factors, e.g., the learning abilities, the quantity and high number of dimensions/features of the data, and the computational resources needed by the AI systems, as also argued in related work (Facchini and Termine 2022)

Grünke (2019) further elaborates on the notion of opacity-3, which they call “epistemic opacity,” through a comparative analysis of the Stockfish (rule-based) and AlphaZero (neural network-based) chess engines. According to the author, the two chess engines exhibit different types of epistemic opacity. Stockfish is epistemically opaque due to the sheer amount of calculations that it does. The number of positions that the engine calculates each second far exceeds human capabilities (up to 60 million positions per second). Consequently, it is impossible for a human to reconstruct and understand step-by-step the process Stockfish follows to make a move. Since this opacity depends on the cognitive capacities and limitations of a human agents, the author calls it “contingent epistemic opacity”. On top of this, AlphaZero exhibits an additional form of opacity. Not only does AlphaZero calculate millions of positions per second, but it also represents features of the game in a way difficult to understand for humans because we do not have equivalent human concepts for some of the features represented in the neural network. This form of opacity, called “fundamental epistemic opacity,” concerns the way in which the neural network models the characteristics of the game. Humans do not understand why AlphaZero makes certain moves because humans and the AI system develop and deploy two different representations of the game.

The distinction between two dimensions of opacity is also supported by Boge (2022). According to the author, deep neural networks (DNNs) are characterised by h-opacity and w-opacity. The former refers to the lack of understanding due to an agent's cognitive limitations and is thus similar to contingent epistemic opacity. The second refers to what is learned by the DNN, i.e., how it models a phenomenon, and thus is similar to fundamental epistemic opacity.

Finally, Héder (2023) identifies additional sources of epistemic opacity. According to the author, AI systems are partially opaque because of the physical complexity of computers. The argument is that the level of investigation of AI systems is their logical model, which, contrary to the commonly held view, is not perfectly embodied. He proposes that any AI that we may encounter is to be investigated as a cyber-physical system in a physical environment. For instance, the thermodynamic processes in computers are fed back into the system as input for randomness, therefore making the comprehension of these complexities necessary for the understanding of the behaviour of the machine. The author points out that, among other factors, self-modification by responding to the environment (via machine learning) plays a pivotal role in raising epistemic opacity. Since most AI systems are characterised by both complexity and self-modification, his arguments apply to most machine learning-based AI systems.

In conclusion, these analyses of the concept of opacity show that it depends both on the purposes and characteristics of its users, and on intrinsic properties of AI systems. The different sources and types of opacity are considered hindering factors in the user's ability to understand a system. Among these factors, the scale and complexity of AI systems play a pivotal role in generating epistemic opacity, as acknowledged in all the analyses considered.

3.2. Transparency and explainability

In the current debate, transparency and explainability often overlap, along with the related concept of interpretability. In this section, we adopt the taxonomy proposed by Lipton (2018) and (Barredo Arrieta et al. 2020) in order to clarify the meaning of these concepts.

According to Lipton (2018), transparency and explainability are aspects of interpretability. Transparency is understood as the opposite of opacity and so "it connotes some sense of understanding the mechanism by which the model works" (Lipton 2018, 12). Transparent models are usually considered understandable by design. These models can be transparent at different levels: at the level of the entire model (simulatability), at the level of individual components (decomposability), and at the level of the training algorithm (algorithmic transparency) (Barredo Arrieta et al. 2020; Lipton 2018). A transparent simulatable model is such that a human should be able, given the input data, to produce the model's output in a reasonable time through the same calculations as performed by the model. A transparent decomposable model is a model whose parts (inputs, parameters, calculations) are such that they can be interpreted and explained. Algorithmic transparency refers to the

learning algorithm rather than the model itself. Algorithmic transparency only requires knowledge of the algorithm and not of the data or learned model, since it just concerns how the algorithm creates the model. According to Barredo Arrieta et al. (2020, 88), “a model is considered to be transparent if by itself it is understandable”. Based on this definition, Barredo Arrieta et al. (2020) and Lipton identify a number of models, e.g., linear/logistic regression, decision trees, Bayesian models etc., that are transparent in at least one of the above senses, i.e., simulatability, decomposability, algorithmic transparency, collectively referred to as “transparent models”.

As far as the concept of explainability is concerned, Lipton (2018) understands it as a form of post-hoc interpretability. This approach involves generating explanations for the outputs of AI systems after they have been produced. Unlike built-in interpretability, where the decision-making process is transparent from the outset, post-hoc interpretability seeks to retrospectively elucidate how a system arrived at its conclusions. It aims to demystify complex AI models by examining their outputs and deducing the contributing factors, thereby providing insights into the system’s operational logic. This approach appears to be equivalent to “local explainability” from the field of explainable AI (Héder 2023). Similarly, according to Barredo Arrieta et al. (2020, 92), explainability techniques “aim at communicating understandable information about how an already developed model produces its predictions for any given input”. Explainability techniques are used when the model is not transparent in any of the senses above. In this case, understanding cannot be achieved by directly examining the model because the model itself is opaque enough to prevent any form of understanding. To overcome this problem, an understanding of the model is sought out through an explanation external to the model that provides additional information about it. Explanations can be provided, for example, in the form of textual explanations, visualisations, saliency maps, and examples (Lipton 2018). Within the post-hoc explainability techniques, the primary distinction is between model-agnostic and model-specific techniques. The former can provide explanations regardless of the model considered, while the latter can be applied only to particular kinds of models, e.g., deep neural networks.

From this brief discussion, we agree with (Barredo Arrieta et al. 2020, 88) that “*understandability emerges as the most essential concept in XAI*”. Consequently, the concept of understandability is also crucial for clarifying the “transparency-trust thesis”. For this reason, instead of dwelling on more in-depth analysis on the concepts of transparency and explainability, we prefer to go straight to the heart of the matter and confront the concept of understandability. In the next section, we try to shed light on this concept in order to clarify what is meant by “understanding an AI system”.

3.3. Understandability

The conceptual analysis undertaken so far indicates that the challenge in explaining or making AI systems transparent primarily revolves around the issue of understanding such systems. Consequently, this shifts our focus from transparency and

explainability to understandability, prompting an investigation into the meaning of “understanding an AI system”.

Addressing this question, researchers in the field of Explainable AI often draw upon insights from epistemology. For instance, Paez (2019) identifies two distinct categories of understanding applicable to AI: objectual understanding and understanding-why. Objectual understanding involves grasping the relationships within the system, akin to understanding the parts and their interconnections within a whole. This form of understanding parallels that provided by transparent-by-design models (see section 3.2). On the other hand, understanding-why goes beyond this, as it encompasses the ability to engage with counterfactual scenarios and predictions. This level of understanding aligns with the goals of post-hoc explainability techniques.

Additionally, Paez (2019) proposes an alternative dichotomy: mechanical understanding versus functional understanding. Functional understanding focuses on the purposes and functionalities of a system. In contrast, mechanical understanding delves into the specific components, processes, and immediate causal mechanisms within that system. Paez (2019) exemplifies this through the understanding of an alarm clock. One can understand the alarm’s function either through the lens of mechanical understanding – a completed circuit activating the buzzer – or through functional understanding – the clock set to awaken its owner at a designated time.

These different kinds of understanding AI of systems do not resolve the question of the desired depth of understanding we want in the context of AI systems. Paez (2019) posits that within the realm of AI, both objectual understanding and understanding-why are intimately connected, and so we need both. He further argues that mere functional understanding is insufficient, advocating for an integration of mechanical understanding as well.

While we agree with Paez (2019) on this point, it is important to note his omission regarding the specific characteristics that make a system understandable. Asserting that understanding an AI system involves engaging with counterfactual scenarios and predictions raises the critical question of our capability to do so and under which circumstances. This consideration is paramount, as it pertains to our fundamental ability to understand certain AI systems.

We propose that the following characteristics contribute to the understandability of AI systems: linearity, monotonicity, few (and simple) interactions among features, rule-based nature (Molnar 2022). By identifying these characteristics, we can better understand the factors that make AI systems more understandable to humans.

A model is linear if it describes the connection between its inputs and outcomes using a linear, or straight-line, function. This linear relationship is easier to comprehend than a non-linear one, as it represents a simpler kind of connection. Similarly, a model is monotonic if any increase or decrease in an input consistently leads to either an increase or a decrease in the output, depending on the nature of the function. These types of relationships are more straightforward to understand compared to those where changes in inputs and outputs do not follow a regular pattern.

It is also beneficial for a model to have a limited number of simple features that interact in straightforward ways. This simplicity makes it easier to track how the

model is working and what it is trying to represent. Additionally, if the model uses simple inputs and is based on clear rules, like IF-THEN statements, it becomes more understandable.

These characteristics are emphasised for AI systems due to their role in enhancing the system’s comprehensibility. However, it’s important to note that as the size and complexity of these AI systems expand, their understandability can diminish, even with linear or monotonic models. Similarly, decision trees start off as straightforward but can lose clarity as they become more intricate. The same applies to rule-based systems.

The main point here is that simple AI systems are easier to understand. This is in line with the observation that the difficulty in understanding AI systems often stems from their scale and complexity. Thus, a simpler AI system is more likely to be understandable, allowing for deeper insights into its functioning and decision-making processes, unlike complex systems with numerous parameters and intricate interactions.

This conclusion is supported by an additional remark. In section 3.1., we found that the opacity of AI systems is largely due to their complexity. Consequently, it makes sense to argue that a non-opaque, i.e. understandable, AI system is a simple system. In the case of simple AI systems, we can have meaningful insights about its internal states, comprehend specific outcomes, and make accurate inferences and predictions about the system’s behaviour. This is not the case for complex, large, and opaque AI systems, which have millions or billions of parameters and non-linear interactions, making it challenging to grasp their inner workings and decision-making processes.

3.4. Trust

The “transparency-trust thesis” implies that transparency and explainability have a positive effect on trust as they counterbalance the negative effect of opacity on trust. However, what is meant by trust in the context of AI is difficult to determine accurately because research in this field uses different definitions of trust. In this section, we provide our understanding of “trust in AI systems” in the context of the “transparency-trust thesis”.

Much of the debate on this topic revolves around the question of whether it makes any sense at all to talk about “trust in AI systems”. Freiman (2023) offers a clear analysis of why this expression is conceptual nonsense according to numerous philosophical accounts of trust. In a nutshell, talking about trustworthy AI systems does not make sense because trustworthiness entails human qualities, such as responsibility, morality, intentionality. Since AI systems lack these human qualities, humans cannot meaningfully trust an AI system. Accordingly, it makes no sense to talk about a “trustworthy AI”. At most, one should talk about “reliable AI” (Freiman 2022, 6). Similarly, Alvarado (2023) argues for a reduced scope of trust in AI. This argument, however, is refuted by other authors, for example Floridi and Sanders (2004), who are comfortable talking about the mind-less morality of artificial agents and therefore find it consistent to talk about “trust in AI systems”.

Both arguments hinge on the assumption that a dependable connection exists between the potential for trustworthy AI systems and the presence of human attributes, like morality and agency, as a prerequisite for these systems to be considered trustworthy. Freiman posits that, due to the lack of such human characteristics in AI systems, discussing the concept of Trustworthy AI is nonsensical. Conversely, Floridi and Sanders argue that, since AI systems do have moral traits, this paves the way for the possibility of them being regarded as trustworthy.

Although it is impossible to completely avoid philosophical assumptions on such subjects, we aim to develop an account of “trust in AI systems” that is less dependent on these assumptions. Specifically, we seek to establish an understanding of trust that does not presuppose any direct relationship between the capacity to place trust and the presence of human qualities in the entity being trusted. In light of this perspective, humans can trust AI systems even in the absence of human-like qualities, and AI systems can be considered “trustworthy” without possessing attributes such as morality, responsibility, or other human-like traits. Consequently, we move away from the positions exemplified respectively by Freiman (2023) and Floridi and Sanders (2004), and develop an account of “trust in AI systems” rooted in studies on “trust in technology” by McKnight et al. (2009).

According to the authors, three meanings of “trust in technology” can be distinguished: trust in specific technology, propensity to trust general technology, and institution-based trust in technology.

Trust in a specific technology is defined as “a willingness to depend on the specific technology in a given situation in which negative consequences are possible” (McKnight et al. 2009, 7). This kind of trust is formed by two components: trusting intention, i.e., willingness to depend on the technology, and trusting beliefs, i.e., the judgement that the technology has desirable attributes, particularly reliability, functionality, and helpfulness. Trusting beliefs are positively related to trusting intention, as individuals are more willing to depend on technology that they believe is trustworthy.

Propensity to trust general technology refers to the willingness to “trust technology across situations and persons” (McKnight et al. 2009, 7). It can take two forms. On the one hand, the propensity to trust in general technology may result from the belief that technology is usually consistent, reliable, functional and helpful. On the other hand, it may be reflected in the belief that positive outcomes will result from the use of technology.

Finally, institution-based trust in technology arises from the belief that, when using technology, “success is likely because of supportive situations and structures” (McKnight et al. 2009, 9). This form of trust depends either on the belief that “success with the specific technology is likely because one feels comfortable or favourable when one uses the general type of technology of which this specific technology is an instance” or when “regardless of the characteristics of the specific technology, one believes structural conditions like guarantees, contracts, support, or other safeguards exist in the general type of technology that make success likely” (McKnight et al. 2009, 9).

Building on this account of “trust in technology”, we define “trust in AI systems” as *the attitude of a person to accept AI decisions or to rely on an AI system to perform a task, accompanied by the belief that the system will make a decision or perform the task in line with the person’s expectations*. This definition is consistent with, and indeed extends, the previous account of trust in technology provided by McKnight et al. (2009). First, it takes into consideration the two components of trust in specific technology, i.e., trusting intention and trusting beliefs. The person’s willingness to rely on an AI system to perform a task represents the trusting intention, while the belief that the system will perform the task or make a decision in line with their expectations represents the trusting belief. On top of this, our definition also implies a distinction between people who actively use AI systems and thus can decide to rely on them, and people who are passive subjects in their relationship with AI systems, and thus can only accept AI decisions. This extension reflects the growing role of AI systems in decision-making processes, and the importance of trust in these systems for individuals who are subject to AI-generated decisions. For example, if we consider the use of an AI system for disease diagnosis, a doctor is an active user because they rely on the AI system to make a diagnosis, while the patient is a passive subject because they do not use the AI system but are affected by its decision. When there is trust in AI systems, both doctor and patient show the intention and belief to rely on the system and accept its decisions respectively. Consequently, when we talk about “trust in AI systems”, it is important to identify who is trusting, because it helps in understanding the different perspectives and expectations of these two groups. Those who actively use AI systems and those who are affected by its decisions may have different requirements, criteria and thresholds for trust, and it is therefore important to take them into account for designing and developing of AI systems aligned with the expectations and needs of both active users and passive subjects.

Secondly, our definition considers “trust in AI systems” as a particular case of trust in general technology, in that it is described as an attitude and is therefore similar to the propensity to trust in general technology. Consequently, when the objective is to build trust in AI systems, one does not start from scratch, but from a pre-existing level of trust in related technologies. This is important to take into account when developing strategies to build trust in AI systems.

Finally, the concept of institution-based trust is also incorporated in our definition, as we acknowledge that the expectation of positive results from the use of technology can be motivated by the belief that there are institutional arrangements in place that make technology trustworthy. Regulations, standards, certifications, market surveillance, and assurance mechanisms can play a crucial role in building trust in AI systems.

Although our definition still entails certain philosophical assumptions, such as trust as an attitude and trust as a function of technology’s properties, we believe it provides a more practical understanding of the concept of trust in the context of the “transparency-trust thesis”. What the proponents of this thesis are concerned about is that the opacity of AI systems hinders the uptake of AI and prevents diffusion of its benefits, which is why they propose transparency and explainability as a solution. User trust, and more broadly public trust, is instrumental in fostering the uptake of

AI. With “trustworthy AI”, users will use artificial intelligence more, and the general public will be more willing to be subjected to AI-based or AI-assisted decisions relevant to their lives.

4. Transparency and explainability don’t build trust in AI systems

After the conceptual clarification of the main terms involved in the “transparency-trust thesis”, we can now turn to verifying its conceptual consistency and validity. In this section, we provide some arguments to support the idea that when the objective is to create trust in AI systems, transparency and explainability are neither necessary nor sufficient and therefore it is not rational to pursue them for this reason alone. This amounts to a rejection of the “transparency-trust thesis”. This does not mean that there are no other reasons to pursue AI transparency, like some epistemic value.

We first elucidate the relationship between the concepts involved in the “transparency-trust thesis”. Proponents of this thesis posit that opacity, transparency, and explainability significantly influence trust in AI systems by affecting their understandability. Opacity, primarily stemming from the scale and complexity of AI systems, diminishes their understandability (Section 3.1). Conversely, transparency and explainability are believed to enhance the understandability of AI systems by diminishing their opacity (Section 3.2). As opacity, transparency, and explainability all pertain to understandability, we concur with Barredo Arrieta et al. (2020, p. 88) that this is the central concept in this discussion. In the “transparency-trust thesis”, understandability acts as a pivotal factor connecting transparency and explainability to trust in AI systems. The argument posits that users are more likely to trust AI systems that are more understandable due to increased transparency and explainability. In contrast, when AI systems are opaque and challenging to grasp, users’ trust may diminish.

Consequently, we examined the concept of understandability (Section 3.3) in the context of AI systems. We first noticed that understandability depends on the user’s background, expertise, cognitive abilities, and familiarity with AI systems. Motivated by this insight, we focused on characteristics of an AI system that promote understandability. We pinpointed qualities such as linearity, monotonicity, limited and straightforward interactions among features, and rule-based nature, ultimately concluding that simplicity is a crucial factor in understanding AI systems. This finding aligns with the notion that the complexity of AI systems, conversely, is a primary source of opacity that hinders understanding.

Finally, we clarified what it means to trust an AI system (Section 3.4). We proposed a working definition based on McKnight et al.’s (2009) account of trust in technology. According to this definition, trust in an AI system is the attitude of a person to accept AI decisions or to rely on an AI system to perform a task, accompanied by the belief that the system will make a decision or perform the task in line with their expectations. In line with some of the empirical studies presented in Section 2, our account of trust in AI systems recognises that accuracy and reliability, as well as

the existence of supportive institutional structures, contribute to building trust in AI systems. Interestingly, understandability plays no role in McKnight et al.'s (2009) account of trust in technology.

Based on these premises, we can now develop a twofold argument against the “transparency-trust thesis”. On the one hand, we reject the assumption that transparency and explainability convey a significant understanding of AI systems. On the other hand, we reject the idea that understandability is a necessary and sufficient element for trust.

In order to support the first side of our argument, we need to clarify what the role of explanations is in promoting the understandability of AI systems. In our account, since understandability is linked to simplicity, of the key issue is identifying a link between explanation and simplicity. In this context, we argue that explanations should work to bridge the gap between the inherent complexity of AI systems and the user's ability to understand their inner workings to the extent that this is simple enough. This is exactly what XAI aims to do. For example, LIME (Ribeiro et al. 2016) is an algorithm designed to explain the predictions made by AI black-boxes. It works by creating a simple linear model that locally approximates the complex AI model to be explained. It does so by perturbing the input features and observing the changes in the output. Based on these observations, LIME learns a sparse linear model that captures the underlying patterns in the AI's decision-making process. Using this linear model, LIME can then generate explanations for the AI's predictions in the form of lists that rank the most important factors contributing to the decision. This simple rank amounts to an explanation of the main elements influencing the underlying complex processes of the original model. Similarly, a visualisation like a saliency map can serve as an explanation by highlighting the most important regions of the input data, such as image pixels, that contribute significantly to the model's outcome. In other words, a saliency map shows which parts of the image have the most influence on the model's decision-making process. These visualisations simplify the understanding of the AI system by focusing on key input areas.

Accordingly, the idea of making AI systems transparent and explainable ultimately amounts to creating simple models for complex systems in a way that the simple model still conveys all important characteristics of the complex system. However, this is not always possible. Several challenges limit the effectiveness of transparency and explainability in conveying significant understanding of AI systems.

First of all, large and complex systems considered transparent by design are not always understandable. For example, the size of decision trees may be such that it is difficult to understand them, even though the operational principle is very simple. As Molnar (2022, 84) writes: “the more terminal nodes and the deeper the tree, the more difficult it becomes to understand the decision rules of a tree”. Consequently, given an input, it may be difficult for a human to understand why a decision tree model produces a certain output.

Secondly, one of the main issues related to post-hoc explainability techniques is that the explanations are not faithful to what the original model computes (Rudin 2019). This is equivalent to saying that the explanations do not actually explain how the model works. Ghassemi et al. (2021) and Rudin (2019) make this point

when analysing saliency maps as an explainability method. For example, highlighting a region of an image only tells where a neural network looks to classify an image. But it says nothing about why it focuses on that region, nor whether this is correct or whether saliency map allows any inductive reasoning about future prediction.

Thirdly, our analysis of opacity-3 (Burrell 2016), fundamental opacity (Grünke 2019), and w-opacity (Boge 2022) shows that there is a limit to the possibility of making complex AI systems simple and understandable. As these studies show, the complexity, the size, the way certain phenomena are modelled by deep neural networks, constitutes an insurmountable obstacle to human understanding. In fact, the best performing models not only are opaque, but they are usually an ensemble of many complex and opaque models. This scale and complexity allow these systems to match or outperform humans in many relevant tasks, and this is where the benefits and risks of AI lie.

Finally, since explainability methods do not reduce the complexity of an AI system, just create a less complex, thus more understandable model of them, they introduce a gap between the explaining model and reality. Due to the inability to reduce complexity of the real system on the one hand and the upper limit of the complexity of a human-understandable model on the other hand, the gap can be very large indeed. The existence of the gap means that several simplifying decisions need to be made while constructing the model, each of them underdetermined by the real system itself. The result is inevitably a less complex but also less accurate model, and we have already seen that the lack of accuracy can undermine trust. In our case it could mean that the trust in the explaining model can come into question.

As for the other side of the argument, even if in the future the development of explainability techniques were to succeed in finally dispelling the black box problem and make any AI system understandable, this would still not be enough to build trust in AI systems. As shown in our account of “trust in AI systems” in section 3.4, trust is a multidimensional construct involving multiple dimensions. However, understandability is not a factor influencing trust in technology. Trusting intentions and beliefs are reinforced when technology has attributes such as consistency, reliability, and helpfulness. Transparency and explainability do not facilitate the attribution of these properties to an AI system. On the contrary, extensive testing and validation, as well as sound institutional arrangements and practices, do.

5. Conclusion

In this paper we have argued against the “transparency–trust thesis” according to which transparency and explainability build trust in AI systems. We first showed that empirical studies on the topic provide conflicting evidence regarding the effect of transparency and explainability on trust. For this reason, we shifted the analysis to a conceptual level, clarifying the main concepts involved in the “transparency–trust thesis”. On this basis, we found that transparency and explainability do not contribute to a complete understanding of an AI system, and that explainability is

not a relevant factor in building trust in AI systems. Accordingly, we concluded by rejecting the “transparency-trust thesis”.

Our conclusions do not imply that the results achieved in the field of explainable AI are invalid, but only that they are not relevant for building trust in AI systems. Anyone wishing to challenge our conclusion would therefore have to demonstrate two claims. The first is that understanding is a constitutive factor of trust in AI systems. The second is that transparency and explainability significantly increase our understanding of large, complex and opaque AI systems. If these issues are explicitly addressed, clarified, and, contrary to our expectations, positively solved, this could give a more solid foundation to the field of Explainable AI.

Meanwhile, our findings suggest that we proceed in a different direction when it comes to building trust in AI systems. On the one hand, as also argued by Ghassemi et al. (2021), rigorous validation could prove the trustworthiness of an AI system even in the absence of an understanding of its inner workings. This validation should be as holistic as possible, involving datasets, algorithms, models, and adopt a socio-technical perspective, considering impacts on individuals and society. On the other hand, institutional arrangements such as regulations, standards, certifications, market surveillance, and assurance may contribute to building institution-based trust in AI systems. In a society where AI systems play an increasingly important role in high-stakes decision-making processes, it is essential that such systems are trustworthy and that both active users and passive subjects trust them. And that they have good reasons to do so.

References

- AI HLEG. “Ethics Guidelines for Trustworthy AI.” Accessed December 31, 2019.
<https://ec.europa.eu/futurium/en/ai-alliance-consultation>
- Alvarado, R. “What kind of trust does AI deserve, if any?” *AI and Ethics* 3, no. 4 (2023): 1169–1183.
<https://doi.org/10.1007/s43681-022-00224-x>
- Andrada, G., Clowes, R. W., and Smart, P. R. “Varieties of transparency: Exploring agency within AI systems.” *AI & SOCIETY* 38, no. 4 (2023): 1321–1331.
<https://doi.org/10.1007/s00146-021-01326-6>
- Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., and Herrera, F. “Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI.” *Information Fusion* 58 (2020): 82–115.
<https://doi.org/10.1016/j.inffus.2019.12.012>
- Bertino, E., Merrill, S., Nesen, A., and Utz, C. “Redefining Data Transparency: A Multidimensional Approach.” *Computer* 52, no. 1 (2019): 16–26.
<https://doi.org/10.1109/MC.2018.2890190>
- Boge, F. J. “Two Dimensions of Opacity and the Deep Learning Predicament.” *Minds and Machines* 32, no. 1(2022): 43–75.
<https://doi.org/10.1007/s11023-021-09569-4>

- Burrell, J. “How the machine ‘thinks’: Understanding opacity in machine learning algorithms.” *Big Data & Society* 3, no. 1 (2016).
<https://doi.org/10.1177/2053951715622512>
- Facchini, A., and Termine, A. “Towards a Taxonomy for the Opacity of AI Systems.” In *Philosophy and Theory of Artificial Intelligence* edited by V. C. Müller, 73–89. Springer International Publishing, 2021.
https://doi.org/10.1007/978-3-031-09153-7_7
- Floridi, L., Cowsls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., and Vayena, E. “AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations.” *Minds and Machines* 28, no. 4 (2018): 689–707.
<https://doi.org/10.1007/s11023-018-9482-5>
- Floridi, L. and Sanders, J. W. “On the Morality of Artificial Agents.” *Minds and Machines* 14, no 3 (2004): 349–379.
<https://doi.org/10.1023/B:MIND.0000035461.63578.9d>
- Freiman, O. “Making sense of the conceptual nonsense ‘trustworthy AI.’” *AI and Ethics* 3 (2023): 1351–1360.
<https://doi.org/10.1007/s43681-022-00241-w>
- Ghassemi, M., Oakden-Rayner, L. and Beam, A. L. “The false hope of current approaches to explainable artificial intelligence in health care.” *The Lancet Digital Health* 3, no. 11 (2021): 745–750.
[https://doi.org/10.1016/S2589-7500\(21\)00208-9](https://doi.org/10.1016/S2589-7500(21)00208-9)
- Grünke, P. “Chess, Artificial Intelligence, and Epistemic Opacity.” *Információs Társadalom* 19, no. 4 (2019): 7-17.
<https://doi.org/10.22503/inftars.XIX.2019.4.1>
- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., and Pedreschi, D. “A Survey of Methods for Explaining Black Box Models.” *ACM Computing Surveys* 51, no. 5 (2019): 1–42.
<https://doi.org/10.1145/3236009>
- Haque, A. B., Islam, A. K. M. N., and Mikalef, P. “Explainable Artificial Intelligence (XAI) from a user perspective: A synthesis of prior literature and problematizing avenues for future research.” *Technological Forecasting and Social Change* 186 (2023).
<https://doi.org/10.1016/j.techfore.2022.122120>
- Héder, M. “The epistemic opacity of autonomous systems and the ethical consequences.” *AI & SOCIETY* 38 (2023): 1819–1827.
<https://doi.org/10.1007/s00146-020-01024-9>
- Kandul, S., Micheli, V., Beck, J., Kneer, M., Burri, T., Fleuret, F., and Christen, M. “Explainable AI: A Review of the Empirical Literature.” *SSRN Scholarly Paper* (2023)
<https://doi.org/10.2139/ssrn.4325219>
- Kartikeya, A. “Examining Correlation Between Trust and Transparency with Explainable Artificial Intelligence.” *Lecture Notes in Networks and Systems* 507 (2022): 353–358.
https://doi.org/10.1007/978-3-031-10464-0_23
- Kizilcec, R. F. “How Much Information? Effects of Transparency on Trust in an Algorithmic Interface.” In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 2390–2395. New York, USA: Association for Computing Machinery 2016.
<https://doi.org/10.1145/2858036.2858402>

-
- Koene, A., Clifton, C., Hatada, Y., Webb, H., and Richardson, R. *A governance framework for algorithmic accountability and transparency*. 2019.
<https://doi.org/10.2861/59990>
- Langer, M., Oster, D., Speith, T., Hermanns, H., Kästner, L., Schmidt, E., Sasing, A., and Baum, K. “What do we want from Explainable Artificial Intelligence (XAI)? – A stakeholder perspective on XAI and a conceptual model guiding interdisciplinary XAI research.” *Artificial Intelligence* 296 (2021).
<https://doi.org/10.1016/j.artint.2021.103473>
- Lipton, Z. C. “The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery.” *Queue* 16, no. 3 (2018): 31–57.
<https://doi.org/10.1145/3236386.3241340>
- Liu, C.-F., Chen, Z.-C., Kuo, S.-C., and Lin, T.-C. “Does AI explainability affect physicians’ intention to use AI?” *International Journal of Medical Informatics* 168 (2022).
<https://doi.org/10.1016/j.ijmedinf.2022.104884>
- Lopez, J., Textor, C., Lancaster, C., Schelble, B., Freeman, G., Zhang, R., McNeese, N., and Pak, R. “The complex relationship of AI ethics and trust in human–AI teaming: Insights from advanced real-world subject matter experts.” *AI and Ethics* 4 (2024): 1213–1233.
<https://doi.org/10.1007/s43681-023-00303-7>
- McKnight, H., Carter, M., and Clay, P. *TRUST IN TECHNOLOGY: DEVELOPMENT OF A SET OF CONSTRUCTS AND MEASURES* (2009).
<https://core.ac.uk/download/pdf/301349124.pdf>
- Molnar, C. *Interpretable Machine Learning. A Guide for Making Black Box Models Explainable*, 2022.
<https://christophm.github.io/interpretable-ml-book/>
- Páez, A. “The Pragmatic Turn in Explainable Artificial Intelligence (XAI).” *Minds & Machines* 29, (2019): 441–459.
- Papenmeier, A., Englebienne, G., and Seifert, C. “How model accuracy and explanation fidelity influence user trust in AI.” In *IJCAI Workshop on Explainable Artificial Intelligence (XAI)* 2019.
- Ribeiro, M. T., Singh, S., and Guestrin, C. “Why Should I Trust You?: Explaining the Predictions of Any Classifier.” In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 1135–1144. New York, USA: Association for Computing Machinery, 2016.
<https://doi.org/10.1145/2939672.2939778>
- Rudin, C. “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead.” *Nature Machine Intelligence* 1, no. 5 (2019): Article 5.
<https://doi.org/10.1038/s42256-019-0048-x>
- Schmidt, P., Biessmann, F., and Teubner, T. “Transparency and trust in artificial intelligence systems.” *Journal of Decision Systems* 29, no. 4 (2020): 260–278.
<https://doi.org/10.1080/12460125.2020.1819094>
- Shin, D. “The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI.” *International Journal of Human Computer Studies* 146 (2021).
<https://doi.org/10.1016/j.ijhcs.2020.102551>

- Wysocki, O., Davies, J. K., Vigo, M., Armstrong, A. C., Landers, D., Lee, R., and Freitas, A. “Assessing the communication gap between AI models and healthcare professionals: Explainability, utility and trust in AI-driven clinical decision-making.” *Artificial Intelligence* 316, (2023).
<https://doi.org/10.1016/j.artint.2022.103839>
- Zednik, C. “Solving the Black Box Problem: A Normative Framework for Explainable Artificial Intelligence.” *Philosophy & Technology* 34, no. 2 (2021): 265–288.
<https://doi.org/10.1007/s13347-019-00382-7>
- Zerilli, J., Bhatt, U., and Weller, A. “How transparency modulates trust in artificial intelligence.” *Patterns* 3, no. 4 (2022).
<https://doi.org/10.1016/j.patter.2022.100455>

Artificial Intelligence in Higher Education: a UTAUT-based Approach to Modelling Student Acceptance

The present research analyses the key factors that determine the acceptance of artificial intelligence and related technologies among university students. This study employs the Unified Theory of Acceptance and Use of Technology (UTAUT) as its theoretical framework. An online cross-sectional survey was conducted among students currently enrolled in higher education. The proposed hypotheses were tested using CB-SEM on the final dataset (n=438). Our results confirmed that performance expectancy and effort expectancy are crucial in forming behavioural intention to use AI, while social influences exert a moderate effect. However, facilitating conditions showed a weak link with both usage intention and actual usage, suggesting that infrastructural factors play a secondary role in shaping technology acceptance. Results imply that the availability of resources alone is insufficient to drive AI adoption and highlight the strategic importance of targeted educational programmes and awareness campaigns to shape students' expectations and attitudes towards AI.

Keywords: *artificial intelligence, AI acceptance, higher education, technology adoption, UTAUT, structural equation modelling*

Acknowledgement

This paper was “supported by the University Research Scholarship Program of the Ministry For Culture and Innovation from the source of the National Research, Development and Innovation Fund”.

Author Information

Kata Horváth, University of Miskolc

<https://orcid.org/0009-0005-3500-7725>

László Molnár, University of Miskolc

<https://orcid.org/0000-0003-1902-6553>

How to cite this article:

Horváth, Kata, László Molnár. “Artificial Intelligence in Higher Education: a UTAUT-based Approach to Modelling Student Acceptance”.

Információs Társadalom XXV, no. 4 (2025): 44–70.

==== <https://dx.doi.org/10.22503/inftars.XXV.2025.4.3> ====

*All materials
published in this journal are licenced
as CC-by-nc-nd 4.0*

1. Introduction

Since its early emergence in the mid-20th century, artificial intelligence (AI) has evolved through cyclical waves of enthusiasm and disappointment, including two major ‘AI winters’ characterised by declining funding and reduced expectations (Russell and Norvig 2022). Despite these discontinuities, recent decades have witnessed an accelerated resurgence and widespread adoption of AI technologies across industries (Grabowski 2024). The concept of artificial intelligence can be defined in multiple ways, encompassing various aspects and dimensions of the technology and its applications. Based on Zhang et al. (2021) AI is a subfield of computer science aimed at developing systems capable of performing tasks that traditionally require human intelligence, such as learning, reasoning, decision making and comprehension. In the present study, AI is conceptualised as a broad category of data-driven algorithmic systems capable of performing these cognitive tasks. Accordingly, our analysis does not focus exclusively on generative AI applications, but rather on AI-based technologies used in higher education more generally, including adaptive learning systems, intelligent tutoring systems, automated assessment tools and generative AI applications. This broader operationalisation allows the investigated determinants of acceptance to reflect individual-level cognitive evaluations and behavioural intentions toward AI as a technological class rather than toward a single application type.

The profound socio-economic impact of AI has been analysed from multiple perspectives (Capraro et al. 2024; Lukianenko and Simakhova 2024; Pintér 2022; Lórinč 2015), consistently highlighting the need for comprehensive social transformation processes (Sosaya-Rodríguez et al. 2024). From a broader innovation perspective, the societal diffusion of new technologies has traditionally been explained by the Innovation Diffusion Theory (Rogers 1983), which emphasises the role of communication channels, social systems and adopter categories in shaping the spread of innovations. While diffusion theory captures macro-level adoption dynamics, the present study focuses on the micro-level psychological mechanisms underlying individual acceptance behaviour. To operationalise the concept of acceptance behaviour, we can draw on numerous well-established models of social psychology (see Ajzen 1985; Ajzen and Fishbein 1980; Bandura 1986; Davis 1986; Deci and Ryan 1985; Triandis 1979; Venkatesh et al. 2012a, 2003, 2016; Venkatesh and Bala 2008), of which the Unified Theory of Acceptance and Use of Technology (UTAUT) is considered to be the most sophisticated construct, providing a concise and thorough depiction of the complexity of factors that have parallel effects on the formation of acceptance behaviour in consumers (Venkatesh et al. 2003).

To establish a solid foundation and define scientifically relevant objectives for our empirical research, we conducted a systematic literature review, analysing scientific papers published between 2018 and 2023. The meta-analysis of 44 relevant studies shows that the topic of AI adoption has been widely studied in various fields such as business, finance and healthcare (Norzelan, Mohamed and Mohamad 2024; Schulz et al. 2023), yet limited research has explored its acceptance in higher education in the given time period. This study aims to address this gap by examining the key factors

influencing university students' adoption of AI-based technologies, as it is a sector that directly impacts societal transformation, influencing the general acceptance and spread of the technology. By fostering AI literacy and advocating ethical use, universities contribute to preparing society for the implications of AI, ensuring that the benefits of this transformation are accessible to all segments of the population (Ferk Savec and Jedrinović 2024). Furthermore, integrating AI into education equips students with relevant digital skills and knowledge, preparing them for evolving labour market demands (Makarenko et al. 2024). Encouraging university students to leverage AI for personal development and to develop practical AI skills is crucial, as AI proficiency is becoming a key requirement across different professions. In light of these considerations, higher education is not only a societal entry point for AI adoption but also emerges as a crucial facilitator in the targeted development of skills that determine the process. Given the expanding role of AI in higher education (Alateyyat and Soltan 2024), understanding student acceptance is becoming critical. Accordingly, this study aims to analyse AI acceptance and its influencing factors among university students. To achieve this objective, an online questionnaire was conducted among Hungarian university students, following the UTAUT model as a theoretical framework. The acquired data was analysed using the CB-SEM method.

This study follows the structure outlined below. Section 2 provides a literature review, summarising previous findings on AI acceptance and explains the theoretical framework used during the research. This section presents our hypotheses and conceptual model as well. Section 3 describes the methodology applied in the survey. Section 4 presents the results of our analysis, which are subsequently discussed in Section 5. Finally, Section 6 summarises the key conclusions of the study.

2. Literature review

2.1. AI in Higher Education

Artificial intelligence is already embedded in higher education, revolutionising traditional academic practices and offering significant advancements in both teaching and institutional management. Based on the available literature, we have identified five main dimensions of AI adoption: (1) personalised education (Jing and Boyi 2023; Murdan and Halkhoree 2024; Naseer et al. 2024; Sajja et al. 2024); (2) support for innovative educational practices (Alrayes, Henari and Ahmed 2024; Saleem et al. 2024); (3) optimisation of institutional operations (Nurhasanah et al. 2024; Téllez, Ortiz and Domínguez 2024; Cholyshkina et al. 2024); (4) facilitation of data-driven managerial decision-making (Fok et al. 2018; Téllez, Ortiz and Domínguez 2024); and (5) transformation of knowledge acquisition and academic research (Dwivedi et al. 2023; Cotton, Cotton and Shipway 2023). Consequently, AI-based solutions guide educational institutions towards more effective, evidence-based practices. Despite the unquestionable advantages that artificial intelligence offers to higher education institutions, its integration poses significant challenges related to ethical considerations (Alrayes, Henari and Ahmed 2024; Héder 2020, 2021), the preservation of educational integrity

(Murdan and Halkhoree 2024) and the digital divide among stakeholders (Murdan and Halkhoree 2024; Téllez, Ortiz and Domínguez 2024). In particular, the emergence of generative AI tools such as large language models has generated extensive debate regarding authorship, academic integrity, and the epistemic role of AI in higher education (Kasneci et al. 2023; Cotton, Cotton and Shipway 2023).

2.2. AI acceptance

Investigating personal or organisational acceptance of emerging technologies has been the focus of several studies. The acceptance of AI and the factors that influence its popularity among individuals have been increasingly discussed in a wide array of publications, covering different sectors and applications from medicine (Schulz et al. 2023), finance (Norzelan, Mohamed and Mohamad 2024) to construction (Na et al. 2022). Artificial intelligence (AI) acceptance can be defined as the willingness and readiness of individuals or organisations to adopt and use AI-based technologies (Kelly, Kaye and Oviedo-Trespalacios 2023). To begin with, perceived usefulness and ease of use are consistently identified as critical determinants of AI acceptance (Aldraiweesh and Alturki 2025; Ibrahim et al. 2024). These factors are often assessed using models like the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (Kelly, Kaye and Oviedo-Trespalacios 2023). Ismatullaev and Kim (2024) concluded that increased transparency and reliability of AI can induce greater trust and acceptance among users. Cornelissen et al. (2022) found that the technical performance and the functionality of AI-based systems can influence user acceptance in the healthcare sector. Apart from technological factors, individual determinants also play a crucial role in the formation of technology adoption. Based on Schiavo, Businaro and Zancanaro (2024), higher AI literacy, i.e. the understanding and ability to use AI technology, reduces anxiety in individuals and fosters a positive attitude towards AI, hence exerting a positive influence on acceptance. Furthermore, certain personality traits, such as openness and an innovative inclination can enhance the perceived ease of use and usefulness of AI, thus prompting acceptance (Hao, Miao and Yan 2021). Trust and anxiety have also proved to be cardinal in the formation of acceptance. The first has a positive effect on AI acceptance, while the latter has a negative impact (Cornelissen et al. 2022; Schiavo, Businaro and Zancanaro 2024). Moreover, organisational factors, such as a supportive culture, participation (Bengel 2020) and a competitive climate (Fousiani et al. 2024) are also vital in the implementation and acceptance of AI technologies. Furthermore, several studies have concluded that social influences, such as social norms, social influence, word of mouth and societal attitudes towards AI can alter the general acceptance of technology (Chen et al. 2024; Fares et al. 2024; Mutlu 2024). Banytė, Lindžiuvienė and Dargytė (2024) proved that privacy and data management concerns can also hinder AI acceptance and emphasised the need for addressing these concerns directly in order to gain trust of users. Finally, emotional and cognitive factors are also significant determinants of AI adoption. In their study, Mutlu (2024) found that the emotions and intrinsic motivations of individuals have

a significant impact on their readiness to use AI-driven technologies and underlined that generally positive responses can enhance acceptance. Based on Chen et al. (2024) and Hao, Miao and Yan (2021), perceived risk associated with AI can have a negative impact on acceptance. Mitigating these risks through guidelines and literacy programs is essential to ensure the success of AI-driven institutional innovations. Nevertheless, the factors influencing AI adoption and the nature of their impact can show a great variance based on the different contexts the tool is being used (Aldraiweesh and Alturki 2025; Nazri et al. 2023).

2.3. Unified Theory of Acceptance and Use of Technology

The present study is based on the Unified Theory of Acceptance and Use of Technology (UTAUT) model. The UTAUT model was developed by integrating and revising eight different models and theories that primarily originate from behavioural sciences. It incorporates certain elements of the Technology Acceptance Model (TAM) while building on the Theory of Reasoned Action (TRA) and the Theory of Planned Behaviour (TPB) as well (Venkatesh et al. 2003). The core variables that UTAUT focuses on are usage intention and actual usage of the technology. Based on the framework, the independent and moderating variables that influence these outcomes can be examined as well. The model developed by Venkatesh et al. (2003) explains consumer technology acceptance through four input variables and one mediator variable, which ultimately reflect in actual usage behaviour. The input variables include individual expectations regarding technology performance (performance expectancy, PE), consumer perceptions of the difficulty associated with using the technology (effort expectancy, EE), the social influence or potential pressure to adopt the technology (social influence, SI), and certain facilitating conditions, such as the infrastructural and other necessary requirements for technology use (facilitating conditions, FC). These factors directly influence usage intention (behavioural intention, BI), which serves as a mediating variable within the model. Additionally, the model establishes a direct relationship between facilitating conditions and the outcome variable, actual usage of technology (use behaviour, UB) (Venkatesh et al. 2003). Among these factors, performance expectancy and effort expectancy have been identified as the strongest predictors of usage intention (Gupta et al. 2019; Simeonova, Bogolyubov and Blagov 2013; Solti 2019). However, in certain cases, the effect of effort expectancy has been found to be non-significant (Sánchez-Holgado and Arcila-Calderón 2024). By contrast, the impact of social influence and facilitating conditions on usage intention appears to be more variable, with their significance showing variance across different research contexts (Bhati, Sharma and Gola 2023; Jewer 2018; Soomro 2019).

Venkatesh et al. (2016) proposed a revised model (UTAUT 2) that supports a more consumer-oriented approach. The model introduces new variables, i.e. hedonic motivation, price value and habit, providing a more complex theoretical framework for the analysis of technology acceptance. Our choice between the two models is guided by several recommendations in the literature. First, the fact that the UTAUT model is a simpler approach compared to UTAUT 2 can be advantageous in contexts

where the analysis aims to uncover fundamental relationships. Tamilmani et al. (2021) highlight that from a parsimony perspective, the application of UTAUT 2 may sometimes be less favourable. Moreover, UTAUT 2 was primarily developed for analysing situations where consumer and consumption-related aspects are prominent. As a result, the model incorporates factors such as hedonic motivation and price value, whose relevance in non-commercial contexts may be questionable. This is supported by the findings of Syamsudin et al. (2018), who investigated public service technologies. Their results indicate that certain model elements introduced to diversify the original framework, such as hedonic motivation, are less applicable to technologies used in public services and often require modification. Given that the higher education sector can also be considered as a non-commercial context, it can be assumed that certain constructs of UTAUT 2 may not be relevant in studies examining technology acceptance in this domain. Furthermore, Or (2023) highlights that in many cases, the application of UTAUT 2 in educational settings does not yield significantly better results than the original model, particularly in explaining the variance between usage intention and actual use. This suggests that in specific educational contexts, the effectiveness of UTAUT is comparable to, or may even surpass that of UTAUT 2. Based on the findings discussed above, we conclude that the original model is more suitable with respect to the objectives of this research.

2.4. Hypotheses and conceptual model

Within the framework of the present research, the hypotheses examined are grounded in the findings of several studies that have investigated the acceptance of various technologies (e.g., e-learning systems, mobile-based learning, interactive whiteboards) used in different educational contexts.

(H1) Performance Expectancy (PE) positively influences university students' Behavioural Intention (BI) to use Artificial Intelligence (AI).

We posit that students' expectations regarding the performance of artificial intelligence have a favourable impact on their intention to use AI technologies. Specifically, the more useful students perceive AI tools and the services they provide, the more open they will be to adopting and utilizing them in their studies. The strong positive relationship between performance expectancy and behavioural intention in the context of higher education technologies has been confirmed by several previous studies (Xue, Rashid and Ouyang 2024). For instance, in the cases of interactive whiteboards (Wong, Teo and Russo 2013) and mobile-based learning (Khan et al. 2022), it has been explicitly demonstrated that performance-related expectations significantly influence both student and teacher intentions to adopt these technologies.

(H2) The Effort Expectancy (EE) positively influences students' Behavioural Intention (BI) to use AI tools.

In general, a positive relationship can be observed between effort expectancy and behavioural intention. However, the strength of this relationship is highly context-dependent. While some studies have identified a moderately significant effect (Ali, Warraich and Butt 2024), others have found a negligible impact of expected effort on behavioural intention (Bacuna and Castro 2023).

(H3) Social Influence (SI) positively affects students' Behavioural Intention (BI) to use Artificial Intelligence.

We hypothesize that students' social environment, including prevailing perceptions and expectations regarding artificial intelligence, positively influences their intention to use AI technologies. This assumption is supported by several studies on technology acceptance in education. However, prior research has reported varying effects of social influence on behavioural intention. For instance, Haripin and Warsono (2024) examined the adoption of an e-learning platform and found a significant relationship between social influence and behavioural intention. In contrast, the study of Bacuna and Castro (2023) on Google Classroom adoption indicated that the effect of social influence on behavioural intention was negligible. These findings suggest that the impact of social networks depends not only on the specific educational context but also on the technology in question.

(H4) Facilitating Conditions (FC) positively influence students' Behavioural Intention (BI) to use Artificial Intelligence.

The availability of necessary resources, knowledge and potential support for using artificial intelligence directly enhances students' intention to adopt AI technologies. This assumption is supported by several studies emphasizing that access to essential resources and assistance plays a crucial role in the acceptance of certain technologies (Bacuna and Castro 2023). However, in some cases, this positive effect is not found to be statistically significant (Ali, Warraich and Butt 2024).

(H5) Facilitating Conditions (FC) Have a Direct Positive Effect on Students' Actual AI Use (UB).

The presence of facilitating conditions supports students' actual use of artificial intelligence. This implies that when students perceive that they have access to the necessary resources and required support, they are more likely to engage in active AI usage. This relationship has been consistently observed in numerous studies (Awwad and Al-Majali 2015; Khechine and Augier 2019). Interestingly, opposing findings have also emerged. For instance, Alblooshi and Hamid (2021) identified a negative relationship between students' use of an e-learning system and facilitating conditions. Their results suggest that merely providing operational and technical resources is insufficient for fostering actual usage behaviour if these conditions do not align with students' specific needs, existing competencies, and prior experiences.

(H6) Students' Behavioural Intention (BI) to use AI positively influences Actual Use (UB).

Numerous studies have confirmed a strong positive relationship between behavioural intention and actual use. For instance, in the context of learning management systems (LMS) during the COVID-19 pandemic, behavioural intention significantly influenced the actual adoption of the technology (Ahmed, Štreimikienė and Štreimikis 2021). Similarly, students' intention to use ICT tools was found to strongly impact their actual usage behaviour (Attuquayefio 2022). This relationship has also been evident in research on online education and e-learning, where studies consistently identified the positive effect of behavioural intention on actual use (Tahir 2023; Zulfakar et al. 2023).

In summary, the theoretical model developed for this study is presented in Figure 1. The model illustrates the measurement models included in the study (manifest and latent variables) as well as the relationships between the various constructs. Latent variables are represented by ovals, while their corresponding manifest variables or measurement indicators are depicted as rectangles. The relationships between the constructs, along with the associated hypotheses, are illustrated through arrows connecting the variables.

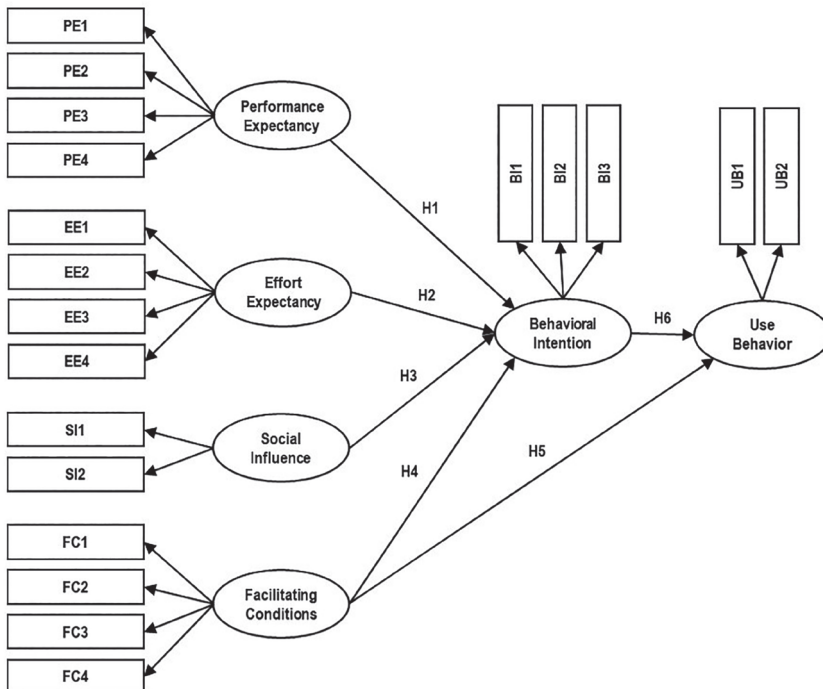


Figure 1. Conceptual model
(Source: edited by the authors)

The proposed model is structured in accordance with the framework of the Unified Theory of Acceptance and Use of Technology (UTAUT) developed by Venkatesh et al. (2003). The dependent variable in the model is actual use (UB), which is directly influenced by the mediating variable (behavioural intention, BI) and the independent variable (facilitating conditions, FC). Four independent variables affect the consumer acceptance of artificial intelligence and the formation of behavioural intention. These include functional elements (performance expectancy, PE, and effort expectancy, EE) as well as contextual factors (social influence, SI, and facilitating conditions, FC). The model is expected to analyse students' intention to use AI along the delineated model variables, identify the mechanisms of influence among the constructs, and determine their significance.

3. Methodology

3.1. Applied Research Method and Sample

This study employed a cross-sectional survey targeting Hungarian students currently enrolled in higher education. Respondents were contacted and invited to participate in the questionnaire-based survey through various social media platforms. Data collection took place between September and November 2024, during which a total of 511 individuals completed the questionnaire. As no predefined sampling frame was available, a non-probability, quota-based sampling technique was applied. Quotas were defined primarily along gender and age dimensions in order to approximate the known demographic composition of Hungarian higher education students. However, due to the social media-based recruitment and the non-random nature of the sampling procedure, the sample cannot be considered statistically representative of the overall population of Hungarian higher education students. Consequently, the findings should be interpreted as applicable to the studied sample, and generalizations to the broader population should be made with caution.

The data cleaning process was carried out in multiple stages. First, responses containing invalid or out-of-range values (e.g., scale responses outside the predefined response categories) were removed (18 records). In addition, responses that failed the embedded attention check item were excluded (2 records). To identify insufficient effort responding, we calculated the within-respondent standard deviation (s) across the scale-based items for each participant. This indicator captures response patterns such as straightlining (very low variance) and excessively inconsistent or erratic responding (very high variance). Cases with $s = 0$ (indicative of complete straightlining) were removed (16 records). In addition, lower and upper threshold values were determined using the Outlier Labeling Rule based on the interquartile range (IQR) of the respondent-level standard deviations. Accordingly, cases with unusually low variability ($s < 0.45$; 10 records) and unusually high variability ($s > 1.55$; 17 records) were excluded from further analyses. Next, we identified logical inconsistencies that could indicate invalid responses. One record was flagged as inconsistent due to a contradiction between the reported age (18–19 years) and the level of

study (PhD), and it was removed (it had already been identified for deletion due to low variance). Duplicate responses were identified using SPSS (3 records), while outliers were analysed using the approach proposed by Olkin and Sampson (2001) approach, employing the Mahalanobis distance test for robust multivariate analysis (7 records). Following the data cleaning process, the final dataset comprised 438 valid responses, which were used for subsequent analyses. Statistical analyses were conducted using SPSS, while the study's hypotheses were tested using covariance-based structural equation modelling (CB-SEM) in SmartPLS 4. The choice of CB-SEM over PLS-SEM was justified by the large sample size, the assumption of multivariate normality, the factor-based modelling approach, and the study's theory-testing focus.

The demographic characteristics of the respondents can be summarized as follows. 51.4% of the respondents were female, while 48.6% were male. All respondents were within the 18-25 age group (mean age: 21 years, standard deviation: 1.68). Regarding place of residence, 21.5% of the respondents lived in rural areas, 18.9% in small towns, 32.6% in cities, and 26.9% in county seats (no responses were received from residents of the capital). In terms of educational programs, the distribution was as follows: 82.6% were enrolled in undergraduate programs, 11.4% in undivided programs, 1.1% in master's programs, 4.1% in higher education vocational training, and 0.7% in specialized postgraduate programs (no responses were received from doctoral students). Regarding prior experience with artificial intelligence tools, 75.6% of the respondents had previously used AI-based tools, while only 24.4% responded negatively to a question regarding their past experience. Given this, we deemed it appropriate to retain the actual usage model construct, as the majority of responses identified this variable as relevant.

3.2. Measurement Methodology

The constructs (latent variables) included in the study, along with the corresponding measurement indicators (manifest variables) and the applied codes, are summarized in Table 1. When defining the individual measurement indicators, we took into account the model elements published by Venkatesh et al. (2003), aiming to create the appropriate higher education context for each indicator. During data collection, respondents were asked to evaluate 19 statements, listed in the following table, using a five-point Likert scale based on their level of agreement. The scale ranged from "1 = Strongly disagree" to "5 = Strongly agree."

| Construct | Measurement item | Code |
|-----------------------------|---|------|
| Performance Expectancy (PE) | Artificial intelligence is an effective aid for my studies. | PE1 |
| | With the help of AI, I can complete tasks more quickly. | PE2 |
| | The use of artificial intelligence can enhance my performance. | PE3 |
| | Using AI increases my chances of achieving better academic results. | PE4 |

| | | |
|------------------------------|--|------|
| | Interactions with artificial intelligence are simple and clear for me. | EE1 |
| Effort Expectancy (EE) | It is easy for me to learn how to use AI and become proficient in its application. | EE2 |
| | I find the use of AI to be straightforward. | EE3 |
| | Mastering the use of AI is not problematic for me. | EE4 |
| | | |
| Social Influence (SI) | Individuals who influence me and my behaviour believe that I should use artificial intelligence. | SI1 |
| | The people who are important to me believe that I should use artificial intelligence. | SI2 |
| Facilitating Conditions (FC) | I have the resources necessary for using artificial intelligence. | FC1 |
| | I possess the knowledge required for using artificial intelligence. | FC2 |
| | Artificial intelligence is not compatible with the other tools I use. | FC3* |
| | I have acquaintances who could assist me if I encounter difficulties with using AI. | FC4 |
| Behavioural Intention (BI) | If AI were available to me, I would definitely use it. | BI1 |
| | If AI were available to me, I believe I would use it. | BI2 |
| | I plan to use artificial intelligence in the future. | BI3 |
| Use Behaviour (UB) | I frequently use artificial intelligence for my studies and completing academic assignments. | UB1 |
| | I feel that it is essential for me to use AI in my studies. | UB2 |

*Negatively worded item was reverse-coded

Table 1. Constructs and measurement items
(Source: edited by the authors)

Prior to the CB-SEM analysis, we validated the structure of the individual constructs through an exploratory factor analysis. As a result of this validation, two indicators (EE1 and FC4) were removed. In the case of EE1, a cross-loading was identified, meaning that the variable was not clearly associated with only one factor but showed relationships with two factors, which significantly undermines the clarity of the factor structure. The reason for removing the FC4 indicator was its low factor loading (0.327). Based on these findings, the final analysis and evaluation involved 17 indicators.

4. Research results

4.1. Convergent and Discriminant Validation

The validation of our model began by evaluating the criteria of convergence and discriminant validity. In this evaluation, we applied the Fornell-Larcker criterion (Fornell and Larcker 1981), which posits that the condition for convergence is that the average variance extracted (AVE) exceeds a value of 0.5. Furthermore, following Hair's (2006) recommendations, it is also required for convergence validity that, in addition to AVE, the standardized factor loadings should exceed 0.5 and the composite reliability (CR) should surpass the threshold of 0.7. As reflected in Table 2, our constructed theoretical model meets all these criteria.

| Construct | Item | M | SD | Factor-loading | Cronbach α | AVE | CR |
|------------------------------|------|------|-------|----------------|-------------------|-------|-------|
| Performance Expectancy (PE) | PE1 | 3.68 | 1.142 | 0.784 | 0.899 | 0.689 | 0.896 |
| | PE2 | 3.71 | 1.241 | 0.891 | | | |
| | PE3 | 3.37 | 1.247 | 0.825 | | | |
| | PE4 | 3.19 | 1.293 | 0.681 | | | |
| Effort Expectancy (EE) | EE2 | 3.73 | 1.101 | 0.823 | 0.892 | 0.737 | 0.892 |
| | EE3 | 3.76 | 1.042 | 0.921 | | | |
| | EE4 | 3.68 | 1.088 | 0.728 | | | |
| Social Influence (SI) | SI1 | 2.22 | 1.127 | 0.821 | 0.844 | 0.733 | 0.847 |
| | SI2 | 2.28 | 1.162 | 0.899 | | | |
| Facilitating Conditions (FC) | FC1 | 3.87 | 1.133 | 0.830 | 0.742 | 0.521 | 0.742 |
| | FC2 | 3.39 | 1.139 | 0.715 | | | |
| | FC3* | 3.38 | 1.304 | 0.586 | | | |
| Behavioural Intention (BI) | BI1 | 3.50 | 1.271 | 1.038 | 0.917 | 0.883 | 0.715 |
| | BI2 | 3.63 | 1.268 | 0.842 | | | |
| | BI3 | 3.76 | 1.267 | 0.592 | | | |
| Use Behaviour (UB) | UB1 | 2.89 | 1.327 | 0.837 | 0.785 | 0.686 | 0.804 |
| | UB2 | 2.31 | 1.267 | 0.688 | | | |

*Negatively worded item was reverse-coded

Table 2. Summary table of averages, standard deviations, convergence, and discriminant validity indicators (Source: edited by the authors)

The discriminant validity of the established model is satisfactory, as none of the correlation values exceeded the 0.85 threshold, which, according to Henseler et al. (2015), would suggest weak discriminant validity. The correlations identified between the constructs are illustrated in Table 3.

| | BI | EE | FC | PE | SI | UB |
|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| BI | | | | | | |
| EE | 0.677 | | | | | |
| FC | 0.559 | 0.767 | | | | |
| PE | 0.778 | 0.607 | 0.469 | | | |
| SI | 0.439 | 0.272 | 0.244 | 0.336 | | |
| UB | 0.654 | 0.473 | 0.394 | 0.758 | 0.463 | |

Table 3. Heterotrait-monotrait (HTMT) ratio matrix
(Source: edited by the authors)

4.2. Reliability

The reliability of the model’s constructs can be assessed using three tests: Cronbach’s alpha (α), average variance extracted (AVE), and composite reliability (CR). A measurement model is considered acceptable if the three criteria mentioned above are met, i.e., $\alpha > 0.5$ (or ideally 0.7), $AVE > 0.5$ (Fornell and Larcker 1981), and $CR > 0.7$ (Malkanathie 2015). As shown in Table 2, for each construct, a Cronbach’s α value of 0.742 or higher is observed, the AVE consistently exceeds 0.521, and the composite reliability (CR) is greater than 0.715 in all cases. These results suggest that the measurement model achieves an optimal level of reliability.

4.3. Model Fit

In addition to the aforementioned, we also examined the absolute and relative fit of the model. For the absolute fit, the chi-square test yielded a value of 771.941 (df=113) with a probability level of 0.000. Furthermore, the χ^2/df ratio was 6.831. The additional absolute fit indices showed the following values: GFI: 0.825; AGFI: 0.763; RMSEA: 0.115; SRMR: 0.256. When evaluating relative fit, the NFI, TLI, and CFI indices were considered. The NFI was 0.85, the TLI was 0.841, and the CFI was 0.868. Overall, based on the recommendations in the literature (Byrne and St 2022; Doğan 2022; Heene et al. 2011; Lide, Mat Daud and Shidrah 2013; Mohammad and Yusoff 2018), the model fit is considered acceptable. Furthermore, it can be stated that the model is suitable for parameter estimation and interpretation.

4.4. Hypothesis Testing and Estimations

The covariance-based structural model was applied with the aim of investigating the previously outlined hypotheses and gaining a deeper understanding of university students' intentions to use artificial intelligence. The results of the hypothesis testing are summarized in Table 4, along with the unstandardized and standardized regression weights measured in the model. In cases where a statistically significant relationship ($p < 0.05$) was confirmed in the predicted direction, the corresponding hypothesis was accepted.

| Hypothesis | Relationship | Regression weights | | | | Standardized regression weights | Result |
|------------|--------------|--------------------|-------|----------|-------|---------------------------------|-----------------|
| | | Est. | S.E. | C.R. (T) | P | | |
| H1 | PE → BI | 0.640 | 0.047 | 13.504 | 0.000 | 0.679 | <i>accepted</i> |
| H2 | EE → BI | 0.308 | 0.056 | 5.501 | 0.000 | 0.306 | <i>accepted</i> |
| H3 | SI → BI | 0.238 | 0.045 | 5.260 | 0.000 | 0.229 | <i>accepted</i> |
| H4 | FC → BI | 0.113 | 0.057 | 1.972 | 0.051 | 0.110 | <i>rejected</i> |
| H5 | FC → UB | 0.132 | 0.077 | 1.725 | 0.085 | 0.101 | <i>rejected</i> |
| H6 | BI → UB | 0.724 | 0.058 | 12.419 | 0.000 | 0.570 | <i>accepted</i> |

Table 4. Unstandardized and Standardized Regression Weights and Hypothesis Testing (Source: edited by the authors)

To ensure the stability of the obtained parameters, we employed the bootstrapping procedure, which generates random samples from the original dataset, allowing us to verify how stable or sensitive the parameter estimates are to slight changes in the data. In the case of hypothesis H5, both the primary and secondary evaluations (bootstrapping procedure) revealed a significantly high p-value, leading to the clear rejection of hypothesis H5. In the primary analysis, hypothesis H4 yielded a p-value very close to the 0.05 threshold, which led to the rejection of hypothesis H4. After applying the bootstrapping procedure, it was confirmed that the relationship outlined in hypothesis H4 was not significant, as the secondary analysis produced a p-value of 0.123 (H5: $p = 0.206$, while H1; H2; H3; H6: $p = 0.000$).

Figure 2 displays the standardized estimates and factor weights, illustrating the relationships between the model constructs and the observed indicators, as well as the model's correction factors.

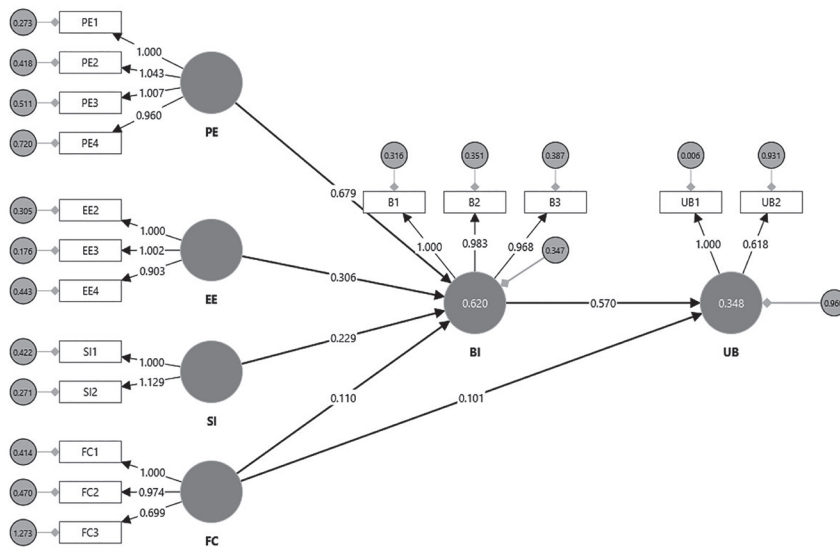


Figure 2. Model results
(Source: edited by the authors)

5. Discussion

In the following sections, we summarise and evaluate the results of our empirical research based on the hypothesis formulated in advance.

First, we suggested that students' performance expectancy regarding artificial intelligence would have a positive impact on their intention to use the technology (H1). This hypothesis was supported ($\beta=0.679$, $p=0.000$), confirming that the anticipated performance benefits of AI play a significant role in shaping university students' usage intentions. In other words, the more useful students perceive a given AI tool to be, the more likely they are to adopt it in their studies. This underscores the importance of explicit communication highlighting the advantages of AI to enhance technology acceptance. The findings are broadly consistent with previous research outcomes. Studies conducted in various educational contexts have consistently identified performance expectancy as a significant predictor of technology usage intention, either in relation to interactive whiteboards or online learning platforms (Wong, Teo and Russo 2013; Yudiantmaja et al. 2022). Furthermore, multiple studies have underscored a positive correlation between performance expectancy and usage intention (Basak, Wotto and Belanger 2018; Xue, Rashid and Ouyang 2024).

Our second hypothesis (H2) proposed a positive relationship between effort expectancy and behavioural intention. The results support this hypothesis ($\beta=0.306$, $p=0.000$). Accordingly, the easier students find operating AI tools or acquiring necessary knowledge to use such technologies, the more likely they are to use them. This finding highlights the importance of user-friendly interfaces and simplified access to AI-based technologies. The positive relationship between effort expectancy

and usage intention has been confirmed in multiple studies examining technologies used in education. For instance, Chao (2019) and Shaya, Madani and Mohebi (2023) investigated the acceptance of mobile-based learning and identified a significant positive correlation between effort expectancy and usage intention. Similar findings have been reported in adaptation studies focusing on ICT tools and virtual reality (VR) headsets (Shen et al. 2017).

Our third hypothesis (H3) proposed a positive relationship between social influence and behavioural intention. Following the analysis, this hypothesis can also be accepted ($\beta=0.229$, $p=0.000$). This suggests that feedback from key reference groups – such as family, friends, and acquaintances – positively influences students' intention to use artificial intelligence for study-related tasks. It also implies that social support and reinforcing peer influence can facilitate AI adoption. The majority of studies in this field reach similar conclusions regarding the relationship between social influence and usage intention in the context of educational technology adoption. For example, a study conducted in Taiwan examined students' behaviours and habits related to an e-learning system and identified a positive relationship between social influence and behavioural intention (Liao, Yu and Yi 2011). This correlation is further reinforced by Haripin and Warsono (2024), who also investigated the acceptance of e-learning systems using the UTAUT model.

The fourth hypothesis (H4) proposed a positive relationship between facilitating conditions and behavioural intention. However, the results did not support this hypothesis ($\beta=0.110$, $p=0.051$). The weak correlation and the p-value approaching the significance threshold indicate that facilitating conditions – such as infrastructure and institutional support – do not have a statistically significant effect on students' intention to use AI. While facilitating conditions are generally considered an important factor in technology adoption, they do not appear to play a decisive role in shaping AI-related usage intentions among university students enrolled in this study. These findings contrast with the prevailing literature, where facilitating conditions have frequently been identified as strong predictors of behavioural intention in educational technology adoption. Studies on e-learning systems in higher education have consistently demonstrated a significant positive relationship between facilitating conditions – such as access to technological infrastructure and human support – and students' willingness to engage with these technologies (Alblooshi and Abdul Hamid 2021; Nain 2021). Nonetheless, there are some empirical findings that support our results, as previous research has also identified instances where facilitating conditions exhibit only a marginal positive effect on technology acceptance. Maita et al. (2018), for instance, found a weak relationship between facilitating conditions and usage intention in the context of academic information systems. Shah et al. (2024) reached similar conclusions. Three potential explanations may account for the negligible relationship observed between facilitating conditions and AI usage intention in this study. One plausible explanation is that the technological infrastructure necessary for AI usage – such as Wi-Fi access, computing resources, and relevant software – may already be widely available at the university. If students take these resources for granted, they are unlikely to perceive them influential in their decision to engage with AI. This aligns with previous research suggesting that

when technological prerequisites are already met, the effect of facilitating conditions on adoption intentions diminishes. Another possible explanation is students' lack of awareness regarding the AI-related resources and support systems provided by the university. If students are unaware of available AI tools, specialised software licenses or dedicated AI learning spaces, they may not factor these resources into their decision-making process. This suggests that insufficient institutional communication about available AI resources could be a contributing factor to the weak relationship observed in this study. A third explanation relates to students' perceived lack of technical preparedness for AI use. If students do not feel sufficiently skilled or trained to use AI effectively, the presence of facilitating conditions alone may not be enough to influence their adoption intention. This perspective is consistent with prior studies highlighting that users' self-efficacy and prior technological experience shape the influence of facilitating conditions on technology acceptance (Venkatesh et al. 2012b; Wang et al. 2020; Al-Marroof et al. 2021).

The fifth hypothesis (H5) proposed a direct positive relationship between facilitating conditions and actual use. Based on our results, the hypothesis was rejected ($\beta=0.101$, $p=0.085$). This suggests that in higher education contexts, the availability of infrastructure and other critical intellectual and technical resources necessary for operating artificial intelligence does not significantly influence students' technology usage intention or their actual usage behaviour. These findings diverge from the general conclusions of previous research in the field. Most studies suggest that facilitating conditions exhibit a significant positive relationship with actual technology use. For instance, prior research has demonstrated that supporting conditions positively influence the adoption of online learning practices (Tahir 2023). However, Handoko (2019) found that among students participating in online learning programmes, facilitating conditions did not have a direct effect on actual use. Similarly, another study by Kuciapski (2016) concluded that facilitating conditions influence actual usage only indirectly, through behavioural intention, rather than via a direct relationship. The weak relationship identified between facilitating conditions and actual use in this study may be explained by several factors. The findings suggest that the formation of actual usage behaviour is primarily driven by motivational factors – such as psychological and social influences – rather than by infrastructural conditions. This implies that students' decisions are largely shaped by their perceptions of the technology's usefulness (performance expectancy, H1) and ease of use (effort expectancy, H2) rather than by the mere availability of supporting conditions. Another possible explanation is that although the necessary technological resources are accessible to students, they may not be optimally utilised. Students may not take full advantage of these resources, especially if they do not perceive immediate and tangible benefits from using them. This scenario is probable when students receive insufficient guidance or practical support in understanding the explicit advantages of AI adoption in their studies. A further factor that may contribute to the weak relationship is the level of independence students maintain from university-provided infrastructure. Many students develop their own learning strategies and may not feel the need to rely on institutional resources. This aligns with the observation that AI tools are increasingly accessible on personal devices such as

laptops and smartphones, reducing students' dependence on university infrastructure. As a result, institutional facilitating conditions may not play a crucial role in shaping students' actual AI usage.

Finally, the sixth hypothesis (H6) proposed a positive relationship between students' intention to use artificial intelligence and their actual AI usage. The results support this hypothesis ($\beta=0.570$, $p=0.000$), reinforcing the notion that university students who are open to AI adoption are more likely to integrate AI into their academic activities. The positive association between behavioural intention and actual technology use in educational contexts has been well documented in prior research. Studies on technology acceptance in online education consistently indicate a strong positive correlation between usage intention and actual adoption (Tahir 2023). Similarly, Awwad and Al-Majali (2015) examined electronic library services at Jordanian universities and concluded that alongside facilitating conditions, usage intention plays a decisive role in determining actual use. These findings align with the broader technology acceptance literature, suggesting that fostering a strong intention to use AI is a critical factor in ensuring its widespread adoption.

6. Conclusions

The empirical research carried out among Hungarian university students has successfully identified the factors that influence the acceptance and adoption of artificial intelligence. In summary, the most significant predictors of artificial intelligence usage are personal perceptions, specifically performance expectancy and effort expectancy. Additionally, social influence contributes to shaping actual usage behaviour, although its impact is comparatively low. By contrast, facilitating conditions, including both human and material infrastructure, did not emerge as key determinants of students' AI usage intention or actual adoption. This finding suggests that while infrastructural availability is important, it does not play a decisive role in driving AI adoption in higher education when students already have access to necessary resources.

Despite its contributions, the study has certain limitations. It focused solely on Hungarian university students, limiting generalisability. Future research should explore diverse educational and cultural contexts. The quantitative approach offered robust insights, but qualitative methods – such as interviews or focus groups – could reveal deeper motivations and barriers to adoption. Moreover, this study examined individual-level factors only, hence future work should also consider institutional elements like faculty support, curriculum design, digital literacy initiatives and formal regulations governing AI use in teaching and assessment. Institutional policies and educators' decisions regarding the permitted or restricted use of AI tools may substantially shape students' actual usage behaviour. Examining such contextual and regulatory factors would provide a more comprehensive understanding of AI adoption in higher education.

The findings offer practical guidance for universities, policymakers, and developers. Higher education institutions should enhance awareness of AI's benefits and

offer training to improve students' perceived ease of use. Encouraging peer collaboration and integrating AI into learning tasks can amplify social influence, which ultimately promotes the formation of acceptable behaviour. Policymakers should support AI literacy, while developers should design intuitive, education-focused AI tools to drive broader adoption.

References

- Ahmed, Rizwan Raheem, Dalia Štreimikienė, and Justas Štreimikis. "THE EXTENDED UTAUT MODEL AND LEARNING MANAGEMENT SYSTEM DURING COVID-19: EVIDENCE FROM PLS-SEM AND CONDITIONAL PROCESS MODELING." *Journal of Business Economics and Management* 23, no. 1 (2021): 82–104.
<https://doi.org/10.3846/jbem.2021.15664>
- Ajzen, Icek. "From Intentions to Actions: A Theory of Planned Behaviour." In *Action Control*. 1985.
- Ajzen, Icek, and Martin Fishbein. *Understanding Attitudes and Predicting Social Behavior*. 1st edn. NJ: Pearson, 1980.
- Alateyyat, Salem, and Mohamed Soltan. "Utilizing Artificial Intelligence in Higher Education: A Systematic Review." In *2024 ASU International Conference in Emerging Technologies for Sustainability and Intelligent Systems (ICETISIS)*, 371–74. Manama, Bahrain: IEEE, 2024.
<https://doi.org/10.1109/ICETISIS61505.2024.10459629>
- Alblooshi, Saeed, and Nor Aziati Binti Abdul Hamid. "The Role of Unified Theory of Acceptance and Use of Technology in E-Learning Adoption in Higher Education Institutions in the UAE." *IBIMA Business Review* (2021): 1–16.
<https://doi.org/10.5171/2021.730690>
- Aldraiweesh, A. A., and U. Alturki. "The Influence of Social Support Theory on AI Acceptance: Examining Educational Support and Perceived Usefulness Using SEM Analysis." *IEEE Access* 13 (2025): 18366–85. Scopus.
<https://doi.org/10.1109/ACCESS.2025.3534099>
- Ali, Irfan, Nosheen Fatima Warraich, and Khadijah Butt. "Acceptance and Use of Artificial Intelligence and AI-Based Applications in Education: A Meta-Analysis and Future Direction." *Information Development* (2024).
<https://doi.org/10.1177/02666669241257206>
- Al-Marouf, Rana Saeed, Khadija Alhumaid, Ahmad Qasim Alhamad, Ahmad Aburayya, and Said Salloum. "User Acceptance of Smart Watch for Medical Purposes: An Empirical Study." *Future Internet* 13, no. 5 (2021): 127.
<https://doi.org/10.3390/fi13050127>
- Alrayes, Amal, Tara Fryad Henari, and Dalal Abdulkarim Ahmed. "ChatGPT in Education – Understanding the Bahraini Academics Perspective." *Electronic Journal of E-Learning* 22, no. 2 (2024): 112–34.
<https://doi.org/10.34190/ejel.22.2.3250>

- Attuquayefio, Samuel NiiBoi. "Information and Communication Technology (ICT) Adoption in Higher Education." *International Journal of Virtual and Personal Learning Environments* 12, no. 1 (2022): 1–21.
<https://doi.org/10.4018/IJVPLE.295309>
- Awwad, Mohammad Sulieman, and Sawsan Mohammad Al-Majali. "Electronic Library Services Acceptance and Use: An Empirical Validation of Unified Theory of Acceptance and Use of Technology." *The Electronic Library* 33, no. 6 (2015): 1100–1120.
<https://doi.org/10.1108/EL-03-2014-0057>
- Bacuna, Rodel D., and Eltimar T. Castro. "Acceptance of Google Classroom Technology among Students at Romblon State University-Cajidiocan Campus Using the Modified UTAUT Model." In *2023 3rd Asian Conference on Innovation in Technology (ASIANCON)*, 1–8. Ravet IN, India: IEEE: 2023.
<https://doi.org/10.1109/ASIANCON58793.2023.10270575>
- Bandura, A. *Social Foundations of Thought and Action: A Social Cognitive Theory*. Vol. 1. Prentice-Hall, 1986.
- Banyté, J., J. Lindžiuviene, and G. Dargytė. "What Factors Determine the Consumer Acceptance of AI-Based Services? The Case of Lithuanian Consumers." (2024): 137–54. Scopus.
https://doi.org/10.1007/978-981-97-1552-7_10
- Basak, Sujit Kumar, Marguerite Wotto, and Paul Belanger. "University Students' M-Learning Adaption Behavioral Factors: A Pilot Study." In *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, 68–73. Vancouver, BC, Canada: IEEE, 2018.
<https://doi.org/10.1109/IEMCON.2018.8615084>
- Bengel, D. *Organizational Acceptance of Artificial Intelligence: Identification of AI Acceptance Factors Tailored to the German Financial Services Sector*. Scopus. (2020).
<https://doi.org/10.1007/978-3-658-30794-3>
- Bhati, Jyoti, Shagun Sharma, and K. R. Gola. "Exploring Factors Behind E-Banking Adoption in India: A UTAUT Model Study." In *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, 2450–54. Coimbatore, India: IEEE, 2023.
<https://doi.org/10.1109/ICACCS57279.2023.10113040>
- Byrne, Barbara, and Cahyono St. *Structural Equation Modeling With AMOS*. 2022.
- Capraro, Valerio, Austin Lentsch, Daron Acemoglu, et al. "The Impact of Generative Artificial Intelligence on Socioeconomic Inequalities and Policy Making." *PNAS Nexus* 3, no. 6 (2024): 191.
<https://doi.org/10.1093/pnasnexus/pgae191>
- Chao, Cheng-Min. "Factors Determining the Behavioral Intention to Use Mobile Learning: An Application and Extension of the UTAUT Model." *Frontiers in Psychology* 10 (2019 July): 1652.
<https://doi.org/10.3389/fpsyg.2019.01652>
- Chen, G., J. Fan, and M. Azam. "Exploring Artificial Intelligence (AI) Chatbots Adoption among Research Scholars Using Unified Theory of Acceptance and Use of Technology (UTAUT)." *Journal of Librarianship and Information Science*, Scopus (2024).
<https://doi.org/10.1177/09610006241269189>
- Cholyshkina, Olga, Andrii Onyshchenko, Volodymyr Kudin, Myroslava Gladka, and Serhii Oleksiienko. "The Use of Artificial Intelligence in Optimising Education Management Processes." *Információs Társadalom* 24, no. 2 (2024): 33.
<https://doi.org/10.22503/inftars.XXIV.2024.2.2>

-
- Cornelissen, L., C. Egger, V. van Beek, L. Williamson, and D. Hommes. "The Drivers of Acceptance of Artificial Intelligence–Powered Care Pathways among Medical Professionals: Web-Based Survey Study." *JMIR Formative Research* 6, no. 6 (2022).
<https://doi.org/10.2196/33368>
- Cotton, Debby, Peter Cotton, and J. Reuben Shipway. "Chatting and Cheating. Ensuring Academic Integrity in the Era of ChatGPT." (2023).
<https://doi.org/10.35542/osf.io/mrz8h>
- Davis, Fred D. *A Technology Acceptance Model for Empirical Testing New End-User Information System: Theory and Results*. Massachusetts Institute of Technology, 1986.
- Deci, Edward L., and Richard M. Ryan. *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer US, 1985.
<https://doi.org/10.1007/978-1-4899-2271-7>
- Doğan, İlkay. "A Simulation Study Comparing Model Fit Measures of Structural Equation Modeling with Multivariate Contaminated Normal Distribution." *Communications in Statistics - Simulation and Computation* 51, no. 5 (2022): 2526–36.
<https://doi.org/10.1080/03610918.2019.1698745>
- Dwivedi, Yogesh K., Nir Kshetri, Laurie Hughes, et al. "Opinion Paper: "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 (2023 August): 102642.
<https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Fares, O. H., Q. Zhu, S. H. Lee, and J. Aversa. "Consumers' Drivers of Generative Pre-Trained Transformer (GPT) Conversational Bot Adoption." In *Revolutionizing the Service Industry With OpenAI Models*. Scopus. (2024).
<https://doi.org/10.4018/979-8-3693-1239-1.ch005>
- Ferk Savec, Vesna, and Sanja Jedrinović. "The Role of AI Implementation in Higher Education in Achieving the Sustainable Development Goals: A Case Study from Slovenia." *Sustainability* 17, no. 1 (2024): 183.
<https://doi.org/10.3390/su17010183>
- Fok, Wilton W. T., Y. S. He, H. H. Au Yeung, et al. "Prediction Model for Students' Future Development by Deep Learning and Tensorflow Artificial Intelligence Engine." In *2018 4th International Conference on Information Management (ICIM)*, 103–6. Oxford, United Kingdom: IEEE, 2018.
<https://doi.org/10.1109/INFOMAN.2018.8392818>
- Fornell, Claes, and David F. Larcker. "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error." *Journal of Marketing Research* 18, no. 1 (1981): 39.
<https://doi.org/10.2307/3151312>
- Fousiani, K., G. Michelakis, P. A. Minnigh, and K. M. M. De Jonge. "Competitive Organizational Climate and Artificial Intelligence (AI) Acceptance: The Moderating Role of Leaders' Power Construal." *Frontiers in Psychology* 15. Scopus. (2024).
<https://doi.org/10.3389/fpsyg.2024.1359164>
- Grabowski, R. "The Future of AI Ain't What It Used To Be." *Design Engineering (Canada)* 69, no. 1 (2024).
- Gupta, Kriti Priya, Rishi Manrai, and Utkarsh Goel. "Factors Influencing Adoption of Payments Banks by Indian Customers: Extending UTAUT with Perceived Credibility." *Journal of Asia*

- Business Studies* 13, no. 2 (2019): 173–95.
<https://doi.org/10.1108/JABS-07-2017-0111>
- Hair, Joseph F., ed. *Multivariate Data Analysis*. 6th ed. Pearson Prentice Hall, 2006.
- Handoko, Bambang Leo. “Application of UTAUT Theory in Higher Education Online Learning.” In *Proceedings of the 2019 10th International Conference on E-Business, Management and Economics*, 259–64. New York, USA: Association for Computing Machinery, 2019.
<https://doi.org/10.1145/3345035.3345047>
- Hao, Z., E. Miao, and M. Yan. “Research on School Principals’ Willingness to Adopt Artificial Intelligence Education and Related Influencing Factors.” (2021): 356–61.
<https://doi.org/10.1109/EITT53287.2021.00076>
- Haripin, and Sony Warsono. “Analysis of the Acceptance of E-Learning Platform (SIDEK-Edu) among High School Students with UTAUT2-TOE Approach: Implications for Digital Learning.” *The International Journal of Information and Learning Technology*, (2024).
<https://doi.org/10.1108/IJILT-03-2024-0051>
- Héder, Mihály. “A Criticism of AI Ethics Guidelines.” *Információs Társadalom* 20, no. 4 (2020): 57.
<https://doi.org/10.22503/inftars.XX.2020.4.5>
- Héder, Mihály. “AI and the Resurrection of Technological Determinism.” *Információs Társadalom* 21, no. 2 (2021): 119.
<https://doi.org/10.22503/inftars.XXI.2021.2.8>
- Heene, Moritz, Sven Hilbert, Clemens Draxler, Matthias Ziegler, and Markus Bühner. “Masking Misfit in Confirmatory Factor Analysis by Increasing Unique Variances: A Cautionary Note on the Usefulness of Cutoff Values of Fit Indices.” *Psychological Methods* 16, no. 3 (2011): 319–36.
<https://doi.org/10.1037/a0024917>
- Henseler, Jörg, Christian Ringle, and Marko Sarstedt. “A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling.” *Journal of the Academy of Marketing Science* 43 (2015): 115–35.
<https://doi.org/10.1007/s11747-014-0403-8>
- Ibrahim, F., J. C. Münscher, M. Daseking, and N. T. Telle. “The Technology Acceptance Model and Adopter Type Analysis in the Context of Artificial Intelligence.” *Frontiers in Artificial Intelligence* 7. (2024).
<https://doi.org/10.3389/frai.2024.1496518>
- Ismatullaev, U. V. U., and S. H. Kim. “Review of the Factors Affecting Acceptance of AI-Infused Systems.” *Human Factors* 66, no. 1 (2024): 126–44.
<https://doi.org/10.1177/00187208211064707>
- Jewer, Jennifer. “Patients’ Intention to Use Online Postings of ED Wait Times: A Modified UTAUT Model.” *International Journal of Medical Informatics* 112 (2018): 34–39.
<https://doi.org/10.1016/j.ijmedinf.2018.01.008>
- Jing, Ren, and He Boyi. “Research on Adaptive Learning Model of Higher Vocational Education Under the Background of Artificial Intelligence.” In *2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, 1–4. Bangalore, India: IEEE, 2023.
<https://doi.org/10.1109/SMARTGENCON60755.2023.10442081>
- Kasneci, Enkelejda, Kathrin Sessler, Stefan Küchemann, et al. “ChatGPT for Good? On Opportunities and Challenges of Large Language Models for Education.” *Learning and*

-
- Individual Differences* 103 (2023): 102274.
<https://doi.org/10.1016/j.lindif.2023.102274>
- Kelly, S., S. A. Kaye, and O. Oviedo-Trespalacios. "What Factors Contribute to the Acceptance of Artificial Intelligence? A Systematic Review." *Telematics and Informatics* 77, 2023.
<https://doi.org/10.1016/j.tele.2022.101925>
- Khan, Farheen Mujeeb, Nidhi Singh, Yuvika Gupta, Jaspreet Kaur, Sandhya Banik, and Sangeeta Gupta. "A Meta-Analysis of Mobile Learning Adoption in Higher Education Based on Unified Theory of Acceptance and Use of Technology 3 (UTAUT3)." *Vision: The Journal of Business Perspective* (2022).
<https://doi.org/10.1177/09722629221101159>
- Khechine, H., and M. Augier. "Adoption of a Social Learning Platform in Higher Education: An Extended UTAUT Model Implementation." (2019): 53–62. Scopus.
<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85072836939&partnerID=40&md5=554f32c1227599895c864157e60123ac>
- Kuciapski, Michał. "Students Acceptance of M-Learning for Higher Education – UTAUT Model Validation." In *Information Systems: Development, Research, Applications, Education*, edited by Stanislaw Wrycza, vol. 264. Lecture Notes in Business Information Processing. Springer International Publishing (2016).
https://doi.org/10.1007/978-3-319-46642-2_11
- Liao, Pei Wen, Chien Yu, and Chin Cheh Yi. "Exploring Effect Factors of E-Learning Behavioral Intention on Cross-Level Analysis." *Advanced Materials Research* 204–210 (2011): 174–77.
<https://doi.org/10.4028/www.scientific.net/AMR.204-210.174>
- Lide, Noor, Nuraihan Mat Daud, and Nor Shidrah. *Interaction Between Writing Apprehension, Motivation, Attitude and Writing Performance: A Structural Equation Modeling Approach*. 21 (2013): 102–8.
<https://doi.org/10.5829/idosi.wasj.2013.21.s1l.2143>
- Lőrincz, András. "Mesterséges Intelligencia Az Egészség És a Jólét Területén." *Információs Társadalom* 15, no. 4 (2015): 54.
<https://doi.org/10.22503/inftars.XV.2015.4.6>
- Lukianenko, Dmytro, and Anastasiia Simakhova. "Artificial Intelligence in the Scientific and Technological Paradigm of Global Economy." *Problemy Ekorożwoju* 19, no. 2 (2024): 55–65.
<https://doi.org/10.35784/preko.6256>
- Maita, Idria, Saide, Richardus Eko Indrajit, and Ade Irmayani. "User Behavior Analysis in Academic Information System Using Unified Theory of Acceptance and Use of Technology (UTAUT)." In *Proceedings of the 2018 1st International Conference on Internet and E-Business*, 223–28. 2018.
<https://doi.org/10.1145/3230348.3230351>
- Makarenko, Oleksandr, Oleksandra Borysenko, Tetiana Horokhivska, Vladyslav Kozub, and Daryna Yaremenko. "Embracing Artificial Intelligence in Education: Shaping the Learning Path for Future Professionals." *Multidisciplinary Science Journal* 6 (2024): 2024ss0720.
<https://doi.org/10.31893/multiscience.2024ss0720>
- Malkanthis, Asoka. *Structural Equation Modeling with AMOS*. 2015.
<https://doi.org/10.13140/RG.2.1.1960.4647>
- Mohammad, Jamilah A. M., and Muhamad S. B. Yusoff. "Psychometric Properties of the Mentor Behaviour Scale in a Sample of Malaysian Medical Students." *Journal of Taibah University*

- Medical Sciences* 13, no. 1 (2018): 58–63.
<https://doi.org/10.1016/j.jtumed.2017.04.005>
- Murdan, Anshu Prakash, and Roshan Halkhoree. “Integration of Artificial Intelligence for Educational Excellence and Innovation in Higher Education Institutions.” In *2024 1st International Conference on Smart Energy Systems and Artificial Intelligence (SESIAI)*, 1–6. Mauritius: IEEE, 2024.
<https://doi.org/10.1109/SESIAI61023.2024.10599402>
- Mutlu, A. Y. “Users Acceptance of Generative Artificial Intelligence Based on ChatBot Use.” 343–48, 2024.
<https://doi.org/10.1109/UBMK63289.2024.10773553>
- Na, S., S. Heo, S. Han, Y. Shin, and Y. Roh. “Acceptance Model of Artificial Intelligence (AI)-Based Technologies in Construction Firms: Applying the Technology Acceptance Model (TAM) in Combination with the Technology–Organisation–Environment (TOE) Framework.” *Buildings* 12, no. 2 (2022). Scopus.
<https://doi.org/10.3390/buildings12020090>
- Nain, Harikishni. “Understanding Students’ Behavioural Intentions To Use Social Media Learning In Education: A Utaut Model Approach.” In *2021 International Conference on Computational Performance Evaluation (ComPE)*, 256–61. 2021.
<https://doi.org/10.1109/ComPE53109.2021.9752450>
- Naseer, Fawad, Muhammad Nasir Khan, Muhammad Tahir, Abdullah Addas, and S. M. Haider Aejaz. “Integrating Deep Learning Techniques for Personalized Learning Pathways in Higher Education.” *Heliyon* 10, no. 11 (2024): e32628.
<https://doi.org/10.1016/j.heliyon.2024.e32628>
- Nazri, I. S. B. M., Z. M. Rodzi, I. N. B. Razali, H. A. Rahman, A. B. A. Rahman, and F. Al-Sharqi. “Unraveling the Factors Influencing the Adoption of Artificial Intelligence (AI) in Education.” 186–93 (2023).
<https://doi.org/10.1109/AiDAS60501.2023.10284698>
- Norzelan, N. A., I. S. Mohamed, and M. Mohamad. “Technology Acceptance of Artificial Intelligence (AI) among Heads of Finance and Accounting Units in the Shared Service Industry.” *Technological Forecasting and Social Change* 198 (2024).
<https://doi.org/10.1016/j.techfore.2023.123022>
- Nurhasanah, Farida, Anjar Sri Ciptorukmi Nugraheni, Bambang Kusharjanta, Roy Ardiansyah, Salim Widono, and Kundharu Saddhono. “A Design of Virtual Classrooms Through AI, ML and DL to Improve the Level of Learning.” In *2024 4th International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 940–45. Greater Noida, India: IEEE, 2024.
<https://doi.org/10.1109/ICACITE60783.2024.10616641>
- Olkin, I., and A. R. Sampson. “Multivariate Analysis: Overview.” In *International Encyclopedia of the Social & Behavioral Sciences* (2001).
<https://doi.org/10.1016/B0-08-043076-7/00472-1>
- Or, Caleb. “Examining Unified Theory of Acceptance and Use of Technology 2 through Meta-Analytic Structural Equation Modelling.” *Journal of Applied Learning & Teaching* 6, no. 2 (2023): 2.
<https://doi.org/10.37074/jalt.2023.6.2.7>

-
- Pintér, Róbert. “Elveszik-e a Gépek a Munkánkat?” *Információs Társadalom* 22, no. 3 (2022): 75.
<https://doi.org/10.22503/inftars.XXII.2022.3.5>
- Rogers, Everett M. *Diffusion of Innovations*. 3rd ed. Free press, 1983.
- Russell, Stuart J., and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Fourth edition, Global edition. With Ming-wei Chang, Jacob Devlin, Anca Dragan, et al. Prentice Hall Series in Artificial Intelligence. Pearson, 2022.
- Sajja, Ramteja, Yusuf Sermet, Muhammed Cikmaz, David Cwiertny, and Ibrahim Demir. “Artificial Intelligence-Enabled Intelligent Assistant for Personalized and Adaptive Learning in Higher Education.” *Information* 15, no. 10 (2024): 596.
<https://doi.org/10.3390/info15100596>
- Saleem, Nudrat, Tabish Mufti, Shahab Saquib Sohail, and Dag Øivind Madsen. “ChatGPT as an Innovative Heutagogical Tool in Medical Education.” *Cogent Education* 11, no. 1 (2024): 2332850.
<https://doi.org/10.1080/2331186X.2024.2332850>
- Sánchez-Holgado, Patricia, and Carlos Arcila-Calderón. “Adoption and Use Factors of Artificial Intelligence and Big Data by Citizens.” *Communication & Society*, 227–46, 2024.
<https://doi.org/10.15581/003.37.2.227-246>
- Schiavo, G., S. Businaro, and M. Zancanaro. “Comprehension, Apprehension, and Acceptance: Understanding the Influence of Literacy and Anxiety on Acceptance of Artificial Intelligence.” *Technology in Society* 77 (2024).
<https://doi.org/10.1016/j.techsoc.2024.102537>
- Schulz, P. J., M. O. Lwin, K. M. Kee, W. W. B. Goh, T. Y. T. Lam, and J. J. Y. Sung. “Modeling the Influence of Attitudes, Trust, and Beliefs on Endoscopists’ Acceptance of Artificial Intelligence Applications in Medical Practice.” *Frontiers in Public Health* 11 (2023).
<https://doi.org/10.3389/fpubh.2023.1301563>
- Shah, Sweety, Neha Mehta, and Anitha Sunil. “Investigation of E-Learning Adoption in Higher Education Based on the Unified Theory of Acceptance and Use of Technology Model.” *E-Learning and Digital Media*, (2024).
<https://doi.org/10.1177/20427530241232493>
- Shaya, N. M., R. A. Madani, and L. M. B. Mohebi. “An Application and Extension of the UTAUT Model Factors Influencing Behavioral Intention to Utilize Mobile Learning in UAE Higher Education.” *Journal of Interactive Learning Research* 34, no. 1 (2023): 153–80.
<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85166415989&partnerID=40&md5=93c24bb756a1885a0b00852012fd3793>
- Shen, Chien-wen, Jung-tsung Ho, Ting-Chang Kuo, and Thai Ha Luong. “Behavioral Intention of Using Virtual Reality in Learning.” In *Proceedings of the 26th International Conference on World Wide Web Companion - WWW ’17 Companion*, 129–37. 2017.
<https://doi.org/10.1145/3041021.3054152>
- Simeonova, B., P. Bogolyubov, and Evgeny Blagov. “Use and Acceptance of Learning Platforms within Universities.” *Proceedings of the European Conference on Knowledge Management, ECKM 2* (2013): 642–51.
- Sobti, Neharika. “Impact of Demonetization on Diffusion of Mobile Payment Service in India: Antecedents of Behavioral Intention and Adoption Using Extended UTAUT Model.” *Journal of Advances in Management Research* 16, no. 4 (2019): 472–97.
<https://doi.org/10.1108/JAMR-09-2018-0086>

- Soomro, Yasir Ali. "Understanding the Adoption of SADAD E-Payments: UTAUT Combined with Religiosity as Moderator." *International Journal of E-Business Research* 15, no. 1 (2019): 55–74.
<https://doi.org/10.4018/IJEER.2019010104>
- Sosaya-Rodríguez, Lilliana Regina, Ronal Manolo Zegarra-Arevalo, Nelson Antonio Angeles-Quinones, and Juan David Jiménez-Vera. *El rol de la universidad en el desarrollo social de la ciencia y la tecnología en el siglo XXI*, 2024.
<https://doi.org/10.5281/ZENODO.11114480>
- Syamsudin, Ruci Meiyanti, Deki Satria, Rinda Wahyuni, and Dana Indra Sensuse. "Exploring Factors Influence Behavioral Intention to Use E-Government Services Using Unified Theory of Acceptance and Use of Technology 2 (UTAUT2)." In *2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, 237–42. Yogyakarta, Indonesia: IEEE, 2018.
<https://doi.org/10.1109/ISRITI.2018.8864474>
- Tahir, M. Mansha. "Students' Behavioural Intention towards Adoption of Online Education: A Study of the Extended UTAUT Model." *Journal of Learning for Development* 10, no. 3 (2023): 392–410.
<https://doi.org/10.56059/jl4d.v10i3.949>
- Tamilmani, Kuttimani, Nripendra P. Rana, Samuel Fosso Wamba, and Rohita Dwivedi. "The Extended Unified Theory of Acceptance and Use of Technology (UTAUT2): A Systematic Literature Review and Theory Evaluation." *International Journal of Information Management* 57 (2021): 102269.
<https://doi.org/10.1016/j.ijinfomgt.2020.102269>
- Téllez, Ramírez Alberto, Fonseca Lina María Ortiz, and Freddy Camilo Triana Domínguez. "Artificial Intelligence in University Administration: An Overview of Its Uses and Applications." *Revista Interamericana de Bibliotecología* 47, no. 2 (2024).
<https://doi.org/10.17533/udea.rib.v47n2e353620>
- Triandis, H. C. "Values, Attitudes, and Interpersonal Behavior." *Nebraska Symp. Motiv.* 27: 195–259, 1979.
- Venkatesh, Morris, Davis, and Davis. "User Acceptance of Information Technology: Toward a Unified View." *MIS Quarterly* 27, no. 3 (2003): 425.
<https://doi.org/10.2307/30036540>
- Venkatesh, Thong, and Xu. "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology." *MIS Quarterly* 36, no. 1 (2012a): 1.
<https://doi.org/10.2307/41410412>
- Venkatesh, Viswanath, and Hillol Bala. "Technology Acceptance Model 3 and a Research Agenda on Interventions." *Decision Sciences* 39, no. 2 (2008): 2.
<https://doi.org/10.1111/j.1540-5915.2008.00192.x>
- Venkatesh, Viswanath, James Thong, and Xin Xu. "Unified Theory of Acceptance and Use of Technology: A Synthesis and the Road Ahead." *Journal of the Association for Information Systems* 17 (2016): 328–76.
<https://doi.org/10.17705/1jais.00428>
- Venkatesh, Viswanath, James Y. L. Thong, and Xin Xu. "Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of

-
- Technology1.” *MIS Quarterly* 36, no. 1 (2012b): 157–78.
<https://doi.org/10.2307/41410412>
- Wang, Hailiang, Da Tao, Na Yu, and Xingda Qu. “Understanding Consumer Acceptance of Healthcare Wearable Devices: An Integrated Model of UTAUT and TTF.” *International Journal of Medical Informatics* 139 (2020): 104156.
<https://doi.org/10.1016/j.ijmedinf.2020.104156>
- Wong, Kung-Teck, Timothy Teo, and Sharon Russo. “Interactive Whiteboard Acceptance: Applicability of the UTAUT Model to Student Teachers.” *The Asia-Pacific Education Researcher* 22, no. 1 (2013): 1–10.
<https://doi.org/10.1007/s40299-012-0001-9>
- Xue, Liangyong, Abdullah Mat Rashid, and Sha Ouyang. “The Unified Theory of Acceptance and Use of Technology (UTAUT) in Higher Education: A Systematic Review.” *Sage Open* 14, no. 1 (2024).
<https://doi.org/10.1177/21582440241229570>
- Yudiatmaja, W. E., Y. Yoserizal, E. Edison, et al. “Adoption of Online Learning in Indonesian Higher Education during the COVID-19 Pandemic.” *Journal of Behavioral Science* 17, no. 2 (2022): 73–89. Scopus.
<https://www.scopus.com/inward/record.uri?eid=2-s2.0-85131453238&partnerID=40&md5=cfe6ccec27554b49cc8d17d785ec19>
- Zhang, Limao, Yue Pan, Xianguo Wu, and Miroslaw J. Skibniewski. “Introduction to Artificial Intelligence.” In *Artificial Intelligence in Construction Engineering and Management*, by Limao Zhang, Yue Pan, Xianguo Wu, and Miroslaw J. Skibniewski, vol. 163. Lecture Notes in Civil Engineering. Springer Singapore, 2021.
https://doi.org/10.1007/978-981-16-2842-9_1
- Zulfakar, Zufara Arneeda, Fitriya Abdul Rahim, Nor Haliza Che Hussain, Azrina Ahmad, Cham Tat-Huei, and Eugene Cheng-Xi Aw. “ZOOM-Ing into a New Pedagogy: Permanent Adoption of Online Teaching and Learning in Private Higher Education Institution in Malaysia.” In *Proceedings of the 2nd International Conference on Emerging Technologies and Intelligent Systems*, edited by Mohammed A. Al-Sharafi, Mostafa Al-Emran, Mohammed Naji Al-Kabi, and Khaled Shaalan, vol. 573. Lecture Notes in Networks and Systems. Springer International Publishing, 2023.
https://doi.org/10.1007/978-3-031-20429-6_39

Could digital competence be one of the most important key competencies of the future?

The study examines the development of digital competence among first-year students of Ludovika National University of Public Service. In response to the growing demands of the labor market and public service, the university introduced a compulsory course aligned with the DigComp 2.1 framework. The training aimed to develop five key areas: information and data management, communication and collaboration, digital content creation, security, and problem solving. The research measured student progress across several semesters using a combination of self-assessment questionnaires, document analysis, and performance-based assessments. The study confirms that targeted, practice-oriented digital competence development can lead to measurable progress in a short period of time. Students evaluated the course structure, teaching methods, and alignment with labor market expectations positively. The case study serves as a replicable model for other higher education institutions and can contribute to the integration of digital literacy into university curricula.

Keywords: *digital competence, higher education, DigComp framework, public administration training, ICT skills development, 21st-century skills*

Acknowledgement

TKP2021-NKTA-51 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-NKTA funding scheme.

Author Information

Márta Katalin Korpics, Ludovika University of Public Service

<https://orcid.org/0009-0004-2971-4565>

Csilla Herendy, Ludovika University of Public Service

<https://orcid.org/0000-0003-3824-0950>

How to cite this article:

Korpics, Katalin Márta, Csilla Herendy. "Could digital competence be one of the most important key competencies of the future?". *Információs Társadalom* XXV, no. 4 (2025): 71–93.

<https://dx.doi.org/10.22503/infars.XXV.2025.4.4>

*All materials
published in this journal are licenced
as CC-by-nc-nd 4.0*

Introduction

The 21st century labor market is undergoing transformations shaped by digitization, automation, globalization, and demographic changes. To remain competitive, it is essential to have transversal key competencies that go beyond professional knowledge and enable individuals to adapt to a rapidly changing environment (OECD 2023). Digital competence encompasses not only technical skills, but also knowledge of information and communication technologies (ICT) and their critical, creative, and ethical use in various contexts of learning, work, social participation, and leisure (Ferrari 2013; Vuorikari et al. 2016, Vuorikari, Kluzer and Punie 2022). It is also a prerequisite for the development of other key competences. Critical thinking and problem solving are closely related to information processing in the digital environment, while adaptability and flexibility can also contribute to security in labor market transitions (Nemeskéri et al. 2016). As the development of digital skills is a lifelong process, it is also closely linked to conscious self-management and career orientation (Carretero, Vuorikari and Punie 2017). The Fourth Industrial Revolution has given rise to new types of jobs that require creative thinking and innovation skills (Schwab 2017). Effective communication and intergenerational cooperation, which are increasingly shaped by digital technologies, are essential for successful participation in the workplace.

Literature review

Higher education institutions are under constant pressure to adapt and develop, which is intensified by the accelerating processes of globalization and digitalization (Budevici-Puiu 2020; Burukina 2021). The information technology revolution has significantly transformed teaching and learning processes, changing students' learning characteristics and expectations of training (Edge et al. 2022). Educational institutions were most strikingly confronted with this change during the Covid-19 pandemic, when digital and online education became a necessity and educational institutions were forced to switch to digital education overnight (Ferri, Grifoni and Guzzo 2020; Ramírez-Hurtado et al. 2021). Although the forced transition brought many challenges, the pandemic also acted as a catalyst, promoting the adoption of new pedagogical and digital methods. In Hungary, following Covid-19, several higher education institutions established teaching methodology and digital competence centers with the aim of developing the digital and pedagogical skills of teachers (Bodnár 2024). At the same time, the challenges generated by these changes are not limited to the technological dimension: the economic environment, funding, student expectations, and social changes (such as increased diversity) all require new responses from higher education (Edge et al. 2022; Sarda, Kasatkina and Vries 2023).

A higher education institution is considered successful if it prepares its students for the labor market. The rapid transformation of the labor market requires skills that even educational institutions often do not have complete, prior knowledge of. At the beginning of the 21st century, competency lists appeared that listed the skills

essential for success in the labor market of the future: global awareness, innovation skills, creativity, media literacy, leadership skills, responsibility, and digital competence (Partnership for 21st Century Skills 2008). Why are 21st-century competencies so important? The reason for this is that there is a competency gap between what is learned in school (whether in public education or higher education) and the competencies needed in practical life (workplace, everyday life) (Vass 2020). This study focuses on digital competence among the aforementioned competencies. For the purposes of this study, it is important to define the concept: „Digital competence means the confident and critical use of information and communication technologies in work, leisure, and communication. The concept encompasses the ability to search for, evaluate, store, create, present, and transfer information, as well as the ability to communicate online and participate in networks“ (Dringó-Horváth et al. 2020, 174). Over the past decade, we have seen the importance of digital competence grow in social, economic, and administrative systems. Digital literacy is not a technological issue, but has become a fundamental social and economic factor that determines the opportunities for individuals and communities to succeed (Budai 2022).

Hungary and the digital divide

The digital divide is a significant challenge in Hungary. It particularly affects older people, those with low levels of education, and residents of rural areas, while the digital skills of younger generations do not always meet international expectations (OECD 2021), as evidenced by declining trends in international assessments, such as PISA results. Both the European Union’s Digital Education Action Plan (European Commission 2021) and Hungary’s Digital Well-being Program 2030 have set the goal of developing digital competencies on a broad scale, but in practice we often encounter fragmented initiatives that do not build on each other. Although access to digital tools has improved in recent years, their use remains uneven, and the education system is unable to keep pace with the rapid pace of digital transformation. Digital skills have become a fundamental asset in the labor market, as the Fourth Industrial Revolution requires new forms of knowledge and skills that are becoming increasingly important as traditional knowledge is becoming increasingly obsolete. The European Union and Hungary have developed several comprehensive strategies, such as the Digital Education Strategy and the Digital Welfare Program, but their effectiveness has been mixed, partly due to slow adaptation and partly due to implementation difficulties. Intervention attempts—such as IKER training, Digital Theme Week, and various device distribution programs—have contributed to the development of digital competencies, but they have typically remained isolated, have not built on each other, and have not offered a comprehensive, systemic solution. The development of digital competence is not only an economic interest but also a social one. A 1% increase in digital literacy can generate an additional HUF 34.7 billion in GDP per year. Failure to develop digital competence, on the other hand, increases social inequalities and economic burdens in the long term (Budai 2022).

The DigComp framework and its significance

Anusca Ferrari's study *DigComp: A Framework for Developing and Understanding Digital Competence in Europe*, commissioned by the European Commission's Joint Research Centre (JRC-IPTS), is widely considered a milestone in the definition of digital competence at the European level. The aim of the document was to provide a common framework for interpreting, developing, and measuring digital competence in the fields of education, the labor market, and social participation. The first version was produced in 2013 (Ferrari 2013), followed by several updates: DigComp 2.0 (2016) contained updated concepts but did not yet define levels (Vuorikari et al. 2016). DigComp 2.1 (2017) introduced eight proficiency levels (Carretero, Vuorikari and Punie 2017), which made it possible to measure skills from basic to advanced levels. DigComp 2.2 (2022) is the latest version and responds to the latest technological and societal challenges, such as the use of artificial intelligence, the fight against disinformation, sustainability, and digital well-being (Vuorikari, Kluzer and Punie 2022).

The framework (DigComp) has identified five main areas of competence:

1. Information and data management – searching for, evaluating, and processing information.
2. Communication and collaboration – responsible use of digital channels, community participation.
3. Digital content creation – creating and editing content, copyright awareness.
4. Safety – data protection, cybersecurity, digital well-being.
5. Problem solving – dealing with technical failures, creative use of new technologies.

The framework breaks down each area into several specific competencies and provides level descriptions to support the development of curricula, training programs, and policy strategies. The material provides support for citizens to use digital technologies and the new systems behind artificial intelligence confidently, critically, and as safely as possible. The material also highlights the new expectations that have emerged in a short period of time in the area of digital competence, which is considered a key competence in the labor market and must be met in most workplaces. Digital readiness and competence are considered competitive factors for 21st-century employees. This is also clearly indicated by the proficiency levels of the reference framework. In addition to other competencies, the development of digital competence has become an important task for higher education, as it is expected that within a few years there will be few workplaces where a high level of digital competence is not required. This is particularly relevant for public administration training, as graduates from these programs will become the operators of e-government.

The working conditions of public service employees are constantly changing, and the digital transition is having a major impact on this sector. The public service needs civil servants who are well equipped to address the challenges of the 21st century.

This places significant responsibility on higher education institutions, as they must apply teaching methods (Tawafak et al. 2020) that equip students with up-to-date skills for dealing with problems in the workplace (Elliott, Bottom and O'Connor 2023; O'Neill 2022; Rubaii 2016). The complex challenges facing the public service are also reflected in the training process, and the responses to these challenges also affect the teaching of individual subjects, requiring a paradigm shift (Benenson et al. 2022; O'Leary 2002). Students must be prepared to deal with new problems in innovative ways in order to perform effectively across various areas of public service (Alford and Brock, 2014; Brock and Alford 2015). The subject of digital competence development is closely linked to the competence development practices and established approaches of students preparing for public administration (Belényesi 2009, 2010; Budai, Csuhai and Tózsza 2023; Horváth 2011).

Developing digital competence requires changes in infrastructure, methodology, and pedagogy on the part of higher education institutions (Farid, Ahmad and Alam 2015; Lepori, Cantoni and Succi 2003; Rogers 2000; Santosa, Batistaa and Marquesa, 2019). One of the fundamental tasks of higher education institutions is to support their students' success in the labor market. In addition to knowledge-based education, the focus has shifted to competence-based education. The presence of competences also contributes to the success of higher education studies, as certain competences are already necessary to complete the studies: 1. attitude (ability to work individually and in a team), 2. learning (passive, active, and future-oriented learning), 3. knowledge (more specialized, profession-specific knowledge and its application), 4. career management skills (Chiu and Chuang 2016). According to another approach, adaptive skills, entrepreneurial skills, teamwork, interpersonal skills, IT skills, technical and field-specific skills, and employability skills are important (Olivier et al. 2014). Since a large percentage of recent graduates typically start work as employees, technical skills and professional knowledge related to the job are most important for them. Higher education competence development should therefore focus primarily on technical skills and practical expertise in undergraduate programs. The possession of ICT competences and the highest possible level of proficiency have become a general and necessary requirement (Budai, Csuhai and Tózsza 2023; Vuorikari, Kluzer and Punie 2022). In Hungary, DigComp is also a reference point in IKER training, the Digital Wellbeing Program, and higher education competency development courses. The example of Ludovika University demonstrates that the framework can be used to develop students' digital skills in a targeted manner and achieve measurable progress in a short period of time (Budai, Csuhai and Tózsza 2023).

In order to develop digital competence, several key questions need to be answered. What do we consider digital competence to be? Is the use of devices consistent with domestic education? How can digital knowledge be measured? (M. Pintér 2019) The situation is more complex than public discourse and the literature suggest (Jenkins 2013; Prensky 2003; Tari 2011). The younger generation may have been born into a digital society and may have relatively higher levels of digital competence, but they do not necessarily have the information and communication skills required by the labor market (Budai, Csuhai and Tózsza 2023; M. Pintér 2019). At the

same time, they really need digital content to complete their studies (Besenyei 2010). The development of ICT skills must begin at secondary school level (Bell, Andreae and Robins 2014) so that a solid foundation can be laid for the development of skills at university.

Digital competence development in higher education – presentation of good practices

Institutional context of good practice

The Ludovika University of Public Service occupies a unique place in the Hungarian higher education system: the institution specializes in training in public administration, public service, and law enforcement. As a result, digital competence is not only a fundamental asset for students' studies, but also for their future professional careers. E-government, digital administration, and data-driven decision-making are areas in which recent graduates must immediately prove themselves (Budai, Csuhai and Tózsza 2023). For this reason, Ludovika University introduced a course on «Digital Competence Development» a few years ago. The course is compulsory for first-year students and is based on the DigComp 2.1 framework. The aim is for students to acquire not only basic user skills but also the higher-level skills that are essential in public administration jobs, such as legally compliant content management, online communication and collaboration, data protection and security, and digital problem solving.

We drew on several sources to present good practices. As background research and empirical results, we took into account the results of a previous study, conducted in the spring semester of the 2022/2023 academic year (Budai, Csuhai and Tózsza 2023). In addition to the questionnaire, we also conducted a document analysis as part of the research, which involved comparing the subject data sheet and the course syllabus. The first survey using the questionnaire we developed took place in the spring semester of the 2023/2024 academic year. The questionnaire was based on a questionnaire used to measure several aspects of the then-prominent paradigm-shifting pedagogical program, thus serving not only to measure competencies but also to measure the methodological and pedagogical outcomes of the course. The course was attended by 177 students, of whom 109 were full-time students and 68 were part-time students. The questionnaire was sent to all participating students. Full-time students filled out the questionnaire in the classroom after the last class, and 62 of the 109 students provided valid responses. Thus, almost 60% of the total number of full-time students responded, which can be considered a high response rate. There were a total of 68 correspondence students on the course, of whom 11 completed the questionnaire. The questions focused on the development of competencies, and students had to assess their own progress in each area. Full-time students had 28 hours (theory) + 28 hours (practice) available for development, while correspondence students had 8 hours (theory) + 8 hours (practice). Due to the low

number of hours in the correspondence course and the low number of responses, we did not evaluate their answers, but only worked with the answers of the full-time students. The questionnaire is administered each semester, and the course syllabus has also been revised to include the use of AI tools and the presentation of VR possibilities among the topics of the foundation course.

The research did not include testing the self-assessment responses. Digital competence was measured directly through classroom assignments, with two practical exercises measuring the application of competencies, and knowledge acquisition was measured through a Moodle test. The effectiveness of classroom tasks is demonstrated by document analysis, which compares the DigComp 2.1 reference framework and the subject syllabus, and also presents tasks aimed at competence development. The results of the research are presented in the form of a case study, which we interpret as a good practice to be followed. There were at least three clearly identifiable reasons for introducing the subject. One was that there were significant differences in the digital competences of the students attending the training. Another was that the training of students graduating from the program places a strong emphasis on public administration informatics and preparation for digital public administration, which they can only successfully complete if they have a solid foundation in digital competencies. The third reason, which we already mentioned in the literature review, is that although we refer to this generation as digital natives, their digital literacy does not extend to the confident use of computer programs (word processing, spreadsheets, database creation, etc.) that are important in their future work.

Research questions

RQ1 Is it possible to achieve demonstrable digital competence development in the competencies defined in the DigComp 2.1. reference framework within the framework of the Digital Competence Development course?¹

RQ2 In which areas of competence can results be achieved during a semester-long course?

RQ3 Do students find the targeted areas of competence useful in relation to their university studies and future workplace?

Results

The questionnaire survey conducted in the spring semester of the 2022/2023 academic year aimed to assess digital competences based on self-reporting. The research addressed three research questions, the first of which was identical to the

¹ The content of the research question is identical to the first question in the previous questionnaire survey for this subject.

first question in our research (Can a significant improvement in digital competence be achieved with a 30-hour theoretical and 30-hour practical university skills development course? The research consisted of an input and an output measurement. The input measurement took place before the start of the course and was completed by 120 students, while the output questionnaire was completed by 58 students. The output measurement took place at the end of the course. Based on the input measurement, students estimated their own digital competence level at 63.3%. They identified information retrieval and data management as their highest rated areas, and digital content creation and problem solving as their weakest areas. After the 30-hour theoretical and 30-hour² practical course, the overall competence level increased by 2.86%. The greatest improvements were in skills (+3.19%) and digital content creation. (+10.65%). The differences between the input and output measurements are shown in Figure 1 (Figure 1). The authors of the study answered the first research question with a clear yes (Budai, Csuhai and Tózsza 2023).

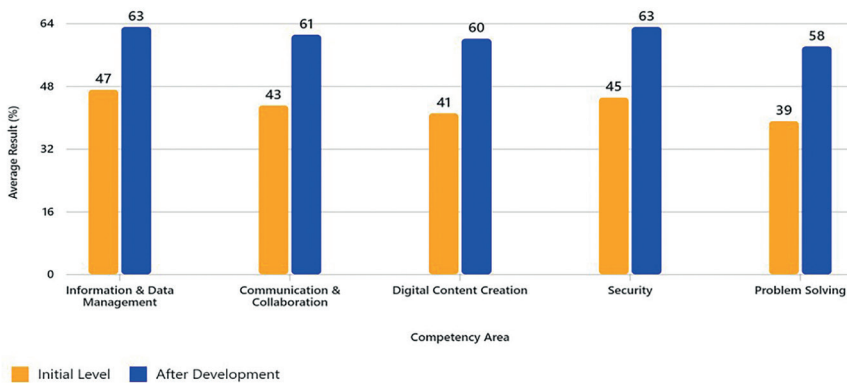


Figure 1: Development of digital competence among students in the spring of 2023 (Source: Own editing based on Budai, Csuhai and Tózsza 2023)

Document analysis

During the document analysis, we compared the DigComp 2.1 reference framework with the course syllabus in order to examine the extent to which the competence areas identified by the reference framework and the course syllabus overlap. The first column of the table contains the competence areas of the reference framework, while the second column indicates their relevance to the course syllabus. The third column lists the areas that were given greater emphasis in the course, which could

² The study mentions 30+30 hours, but given the semester schedule, this means 28+28 hours in practice, as there are no classes scheduled for the last week, which is a make-up week.

be completed for extra credit as homework assignments, and which played a major role in the development of competencies (Appendix 1, Table 1). According to the course syllabus, the aim of the course is to impart skill-level, competency-developing knowledge. In line with the DigComp 2.1 reference framework, it aims to develop proficient digital literacy, complemented by an introduction to the basics of e-administration (public administration and commerce).

The course focused on competence development through classroom assignments, exercises, and homework tasks. There were five assignments related to the classroom tasks, which could earn extra points. The first assignment was to create a resume, which required both text and image editing skills. The resume had to list the 21 competency areas that the students claimed to already possess. The second task was to prepare a 2-page initial text, which required literature research and data collection and had to be formatted according to the specified criteria. The two tasks were related to competence areas 1 (Information and Data Management) and 3 (Digital Content). The third task was to prepare a newsletter, which could be worked on in pairs or trios, on a shared online collaboration platform (Canva). This task was related to competency area 2 (Communication and collaboration). The fourth task was to examine and analyze the website of a public administration organization, which was related to areas 4 (Security) and 5 (Problem Solving). The fifth task was to create a Kahoot test, which tested the topics covered in the course lectures. This task covered many areas and, in addition to testing knowledge, also developed online collaboration (competence area 3). The Kahoot test was completed by students in pairs or trios. The Kahoot test always took place in the first few minutes of the practical lesson.

The findings indicate that development took place in all five competency areas. Based on the instructor's decision, the greatest emphasis was placed on those areas that students will need during their studies and for writing their theses (information and data management, communication, and collaboration). Naturally, development also took place in the other three areas, as the optional homework assignments, which could earn extra points during the semester, were related to all five areas of competence.

Questionnaire Survey

The questionnaire was based on the one used for the university's training development,³ supplemented with a section on digital competence areas. The questionnaire asked about the effectiveness of the course, learning organisation, teaching methods, and progress in the digital competence areas. The questionnaire asked about the structure of the curriculum, perceptions of the course content, opinions on the support of learning methods, the development of digital competences and proportional accountability, aligned with the research questions. Respondents gave a score close to four out of five on the structure of the course content. The course

³ The questionnaire was prepared and evaluated by the staff of the Creative Learning Office.

content itself was rated slightly lower than anticipated (3.79) and the methods used were also rated below four (3.69). For development in the use of digital tools and platforms, students gave a score of 3.66. The highest score was given to the relationship between accountability and course content (4.18) (Figure 2). When measured on a scale of four, student perceptions were basically close to a four, indicating that changing the course content and methods would be worthwhile in terms of their usefulness.

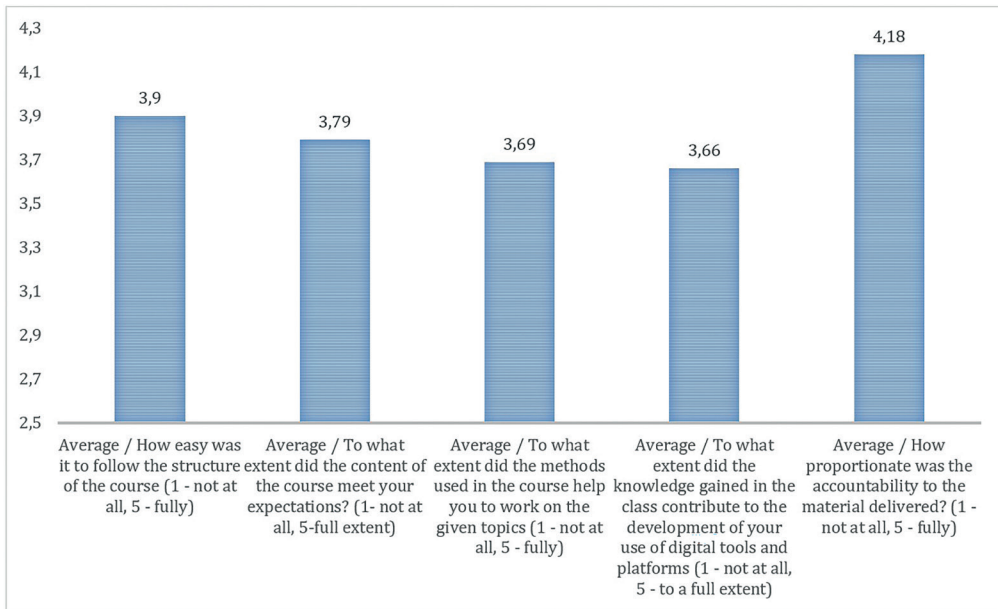


Figure 2: Average scores for questions on the effectiveness of full-time courses (own editing)

The development of digital competence per domain was measured as a percentage. Of the 21 competency areas, we selected those that were developed during the course. For each competency area, we asked what percentage of respondents perceived improvement in that area: 81% of students indicated that they had improved in the area of *digital content editing*, 79% in the area of *copyright and licensing*. *Filtering and evaluating information, browsing, searching and creating digital content* also scored highly (77%). *Creative use of digital technologies* also scored relatively highly (73%), as did *managing digital content* (71%). *Performance evaluation skills* were 10% below 70% (69%). *Digital collaboration* scored well above 60%, and *communication and digital interaction* also scored 66%. *Other skills* above 60% were *solving technical problems* (65%), *protecting personal data and privacy* (63%) and *managing digital identity* (61%). The lowest scores were for *programming* (29%), *protecting the environment* (32%) and *protecting health and well-being* (44%) (Figure 3). The latter did not actually feature in the 14 sessions of the course.

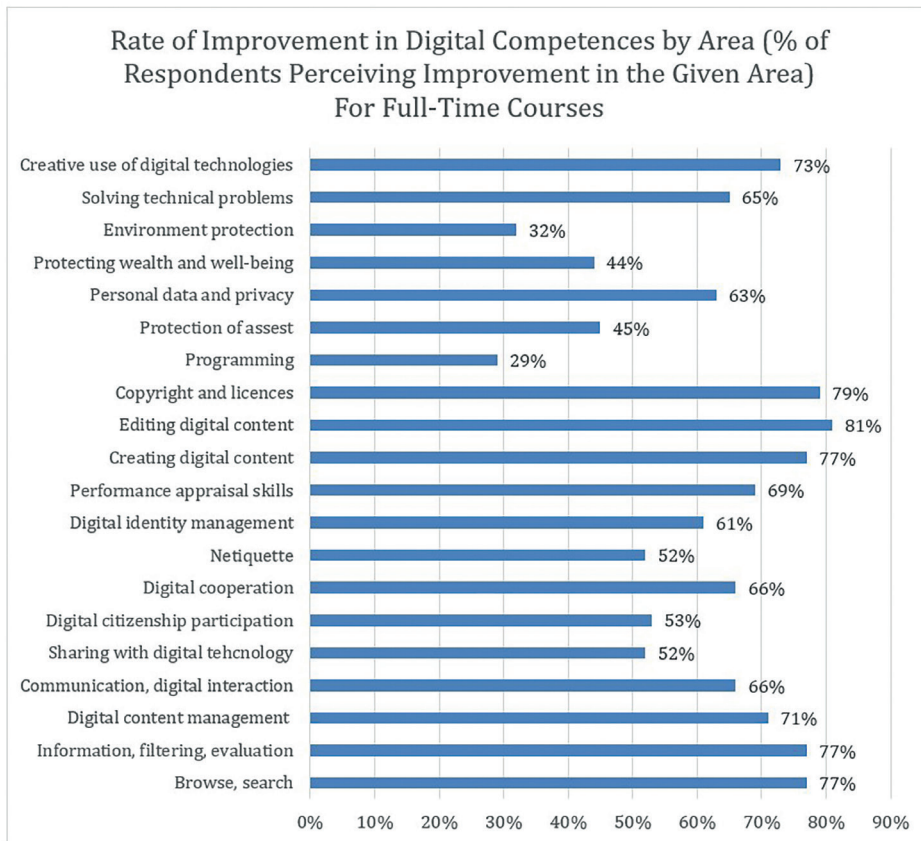


Figure 3: Rate of improvement in digital competences per area (own editing)

The questionnaire also included open-ended questions, which we coded and summarized based on frequency. Among the responses to the questions, we highlight those in which students referred to their achievements in the area of competence: for the positive factors in the lessons, the helpful teacher’s attitude was preceded by a sense of progress and usefulness; and the good classroom atmosphere and extra credit system were also important. Positive factors were, for example, „I learnt how to make a poster and write a CV“, „It was good to have lots of homework for extra credit“ or „The class was very good“, or that „I was happy to have a proper CV that I can submit anywhere and I think it’s important that we learned how to handle references“.

Digital Competence Development at Ludovika University: A Comparison of Two Questionnaire Surveys

We compared the results of the two surveys using the same criteria and examined how similar and different the results were when the course was assessed at two different points in time (Appendix. Table 2). The 2023 study used quantitative methods

to examine changes in competency levels, taking demographic factors into account, while the 2024 study used a qualitative approach, employing document analysis and questionnaire-based data collection. Students evaluated the course on a four-point scale, with the highest score attributed to the proportionality of the curriculum and examination (4.18) and the lowest to the development of digital tools and platform use (3.66). According to the document analysis, there was improvement in all areas of competence, particularly in digital content editing (81%), copyright and licensing (79%), and information filtering, browsing, and content creation (77%). There was a significant improvement in the creative use of digital technologies (73%) and content management (71%), as well as in the area of communication and collaboration. The questionnaire survey confirmed that a large proportion of students experienced improvement in the areas of focus, so overall, both studies provided a comprehensive picture of the opportunities and challenges of digital competence development at Ludovika University.

Further measurements

Measurements taken in the fall of 2024 confirmed that the course resulted in significant digital competency development in a short period of time. Students' self-assessments improved in all dimensions examined, particularly in the areas of digital content creation, information retrieval, data security, and problem solving. The degree of improvement was typically 10 to 20 percentage points and was most pronounced among those who started with weaker competencies. „I won't start here“ category has almost completely disappeared, with the proportion of independent solutions increasing by 10 to 20 percentage points. The greatest practical value was seen among students who were initially weaker. Persistent weaknesses were observed in managing social media privacy settings and complex document formatting (Figure 4.).

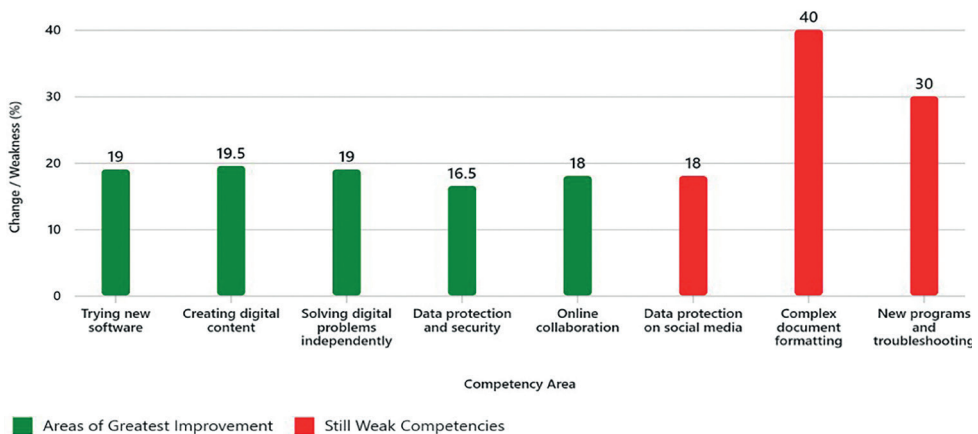


Figure 4: Areas showing improvement and areas that remain weak (own editing)

The greatest improvement was in the areas of trying out new software, creating digital content, problem solving, and data security skills. Basic competencies (file management, searching, email) were already strong, but measurable progress was seen here as well (Figure 5).

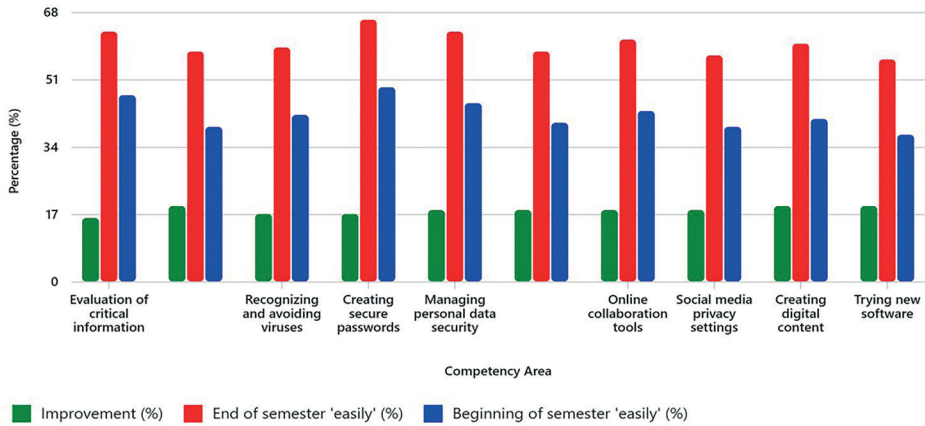


Figure 5. Areas of competence showing the greatest improvement (own editing)

The course brought tangible improvements in the short term, especially among the less confident students, who made the most progress. This is the greatest value of the program

Discussion

Several European Union documents demonstrate that key competences need to be addressed in the field of education, and in particular in higher education. As many aspects of life, including citizenship and learning, become increasingly digitised, digital literacy is highly valued and has become an indispensable prerequisite for navigating and thriving in the knowledge society.⁴ Accordingly, the European Commission amended the DigComp 2.1 reference framework published in 2017 (Carretero, Vuorikari and Punie 2017) in 2022 (Vuorikari, Kluzer and Punie 2022). The new reference framework now includes additional knowledge, skills and attitudes elements. It also indicates how rapidly this area is changing and how new expectations in the area of digital competence are challenging in the workplace.

Higher education must also respond to global and labour market expectations. To prepare future workers, the specialised higher education institution in this study

⁴ European Skills, Competences, Qualifications and Occupations.

is micro-educating its graduates in public administration. Indeed, one of the specific specialisations in the study (BA in Public Administration Management) is public administration informatics. The students graduating from this course are expected to become the future operators of eGovernment, thus responsibility is laid on the training institution. The working conditions of civil servants are significantly affected by the digital transformation. Preparing the right workforce is a key responsibility for higher education institutions, using well-chosen teaching methods (Tawafak et al. 2020) that will equip students with up-to-date competences to deal with the challenges and expectations of the workplace (Elliott, Bottom and O'Connor 2023; O'Neill 2022; Rubaii 2016). Responses to the challenges in public service must already be reflected in the preparation of students, which necessitates a paradigm shift in the teaching-learning process (Benenson et al. 2022; O'Leary 2002). This was also pointed out in the comparative analysis of the competences formulated by the Framework of Reference (Carretero, Vuorikari and Punie 2017) and the subject matter and course content. Moreover, the development of course content and the adaptive introduction of new methodological solutions require continuous attention to enable higher education to meet the ever changing expectations and needs of students (Budai, Csuhai and Tózsá 2023).

Placing the need for competence development within the context of the continuous development of information and technological tools—one of the most rapidly changing areas of globalization—the results of our research can be interpreted as a response to the pressures of change and development that affect higher education institutions. The development of the course content is a good example of learning outcomes-based education (Fryer, McKinney and Trimble 2006; Biggs and Tang 2007), but it also responded well to the development of the competency areas articulated in the Digcomp 2.1 framework. In the case of the degree course we are examining here, it can certainly be said that it is necessary to prepare students for the labour market. Global awareness, innovation, creativity, media literacy and responsibility, which are among the 21st century competences expected of young workers, were all developed on the course. Furthermore, students require these 21 areas of competence to succeed in their studies. The effective implementation of digital competence development requires infrastructural, methodological and pedagogical changes on the part of educational institutions (Farid, Ahmad and Alam 2015; Batistaa and Marquesa 2019; Lepori, Lorenzo and Succi 2003; Rogers 2000). The different levels of knowledge and skills of students also needs to be taken into account (Edge et al. 2022; Sarda, Kasatkina and Vries 2023) which could be compensated for by the introduction of optional but credit-bearing homework assignments. The latter scoring system increased students' academic achievement (Bean 1980; Spady 1971; Kerby 2015). The development of course content here also took into account the development of competences that were not intended to support successful placement in the labour market but rather to support the proper completion of university studies. These developed students' attitudes (through individual and teamwork participation), their ability to learn (passive, active and future-oriented learning), certain knowledge elements (specific, profession-specific knowledge) and career management skills (Chiu-Chang 2016). The latter was contributed to by the

task of preparing digital CVs, as well as the creation of online newsletters and the management and filtering of digital content. Student feedback clearly showed that a high percentage of respondents had improved in this area.

The introduction of the course was necessary because teachers felt that students in the new intakes were becoming decreasingly familiar with computers and computer programs. This is in contrast with what has been reported in the public discourse and literature on the media literacy of digital natives (Jenkins 2013; Prensky 2003). It seems claims about *digital natives* cannot be disputed, but care must be taken in university education to ensure that these young people are able to use not only the tools and applications that they are familiar and comfortable with, but also those that will be expected of them later on, whether as a part of their studies or in the workplace (Budai et al. 2023). The development of ICT competences needs to begin in secondary school, so that the development of competences at university can start with a solid foundation (Bell, Andreae and Robins 2014).

Summary

The study examines the importance of developing digital competence in higher education, with a particular focus on public service training. Due to the demands of the 21st-century labor market and the Fourth Industrial Revolution, digital competence, which includes information management, communication, content creation, security, and problem solving—has become a key competence. The research was conducted among first-year students of the National University of Public Service as part of a compulsory course based on the DigComp 2.1 framework. The study used self-assessment, document analysis, and performance-based tasks. The results confirm that targeted, practice-oriented development can achieve measurable progress in a short period of time, especially in the areas of digital content creation and creative technology use. According to student feedback, the structure and methods of the course are in line with the competence development principles outlined in the literature and meet the needs of the labor market. The study emphasizes that digital competence is not only a technological but also a social and economic factor, the development of which is essential for those preparing for a career in public service. The good practice described can be adapted to other higher education institutions, contributing to the integration of 21st-century skills into university education.

References

- Alford, John, and Janine Brock. "Interactive education in public administration (1): The role of teaching 'objects'." *Teaching Public Administration* 32, no. 2 (2014): 144–157.
<https://doi.org/10.1177/0144739413515491>
- Bean, John P. "Dropouts and turnover: The synthesis and test of a causal model of student attrition." *Research in Higher Education* 12 no. 2 (1980): 155–187.
<https://doi.org/10.1007/bf00976194>

-
- Lepori, Benedetto, Lorenzo Cantoni, and Chiara Succi. "The introduction of e-learning in European universities: models and strategies." In *Digitaler Campus Vom Medienprojekt zum nachhaltigen Medieneinsatz in der Hochschule*, edited by Michael Kerres and Britta Voß, 74–83. Münster: Waxmann, 2003.
- Benenson, Jodi, Tara Kolar Bryan, Carol Ebdon, Theresa Glanz, James Harrold, Thomas Jamieson, and Njoki Mwarumba. "Uncertainty and unrest: A Collaborative Pedagogical Response to Pandemics, Protests, and Policy." *Journal of Public Affairs Education* 28, no. 3 (2022): 269–289.
- Brock, Jonathan, and John Alford. "Interactive Education in Public Administration (2): Strategies for Teachers." *Teaching Public Administration* 33, no. 1 (2015): 6–21.
<https://doi.org/10.1177/0144739414521112>
- Bell, Tim, Peter Andreae, and Anthony Robins. "A Case Study of the Introduction of Computer Science in NZ Schools." *ACM Transactions on Computing Education* 14, no. 2 (2014): article 10.
<https://doi.org/10.1145/2602485>
- Belényesi, Emese. *E-közigazgatás a magyar felsőoktatásban és továbbképzésben* [E-government in Hungarian higher and further education]. Budapest: E-Government Alapítvány a közigazgatás modernizációjáért, 2009.
- Belényesi, Emese. "Kompetenciafejlesztés a közigazgatásban." [Skills development in public administration.] *Jelenkori Társadalmi és Gazdasági Folyamatok* 5, no. 1-2 (2010): 95–100.
- Besenyey, István. "Az unalom lázadása – digitális nemzedék az egyetemen." In *Korszerű felsőoktatási pedagógiai módszerek, törekvések* [Modern pedagogical methods and efforts in higher education], NFKK füzetek 5, edited by István Besenyey, István Perjés, and József Temesi, 40-48. Budapest: Budapesti Corvinus Egyetem NFKK, 2010.
- Biggs, John, and Catherine Tang. *Teaching for Quality Learning at University*. Maidenhead: McGraw-Hill/Open University Press, 2007.
- Bodnár, Éva. "A módszertani központok szerepe a hazai felsőoktatásban: igény vagy kényszer." [The role of methodological centres in Hungary: need or constraint]. *Educatio* 33, no. 2 (2024): 210–221.
<https://doi.org/10.1556/2063.33.2024.2.9>
- Budai, Balázs. „A Digitális kompetencia növekvő szerepe.” *Pro Publico Bono-Magyar Közigazgatás* 10 (2). (2022): 30–59.
<https://doi.org/10.32575/ppb.2022.2.2>
- Budai, Balázs Benjámin, Sándor Csuha and István Tózsá. "Digital Competence Development in Public Administration Higher Education." *Sustainability* 15, no. 16 (2023): 12462.
<https://doi.org/10.3390/su151612462>
- Budevici-Puiu, Liliana. "The Necessity of Change and Development of the Higher Education Institution in the Age of Globalisation." *Revista Romaneasca pentru Educatie Multidimensionala* 12, no. 1 (2020): 350-356.
<https://doi.org/10.18662/rrem/218>
- Burukina, Olga. "University Internal Communication in the Digital Era: Finding a Systemic Approach." *Journal of Emerging Trends in Marketing and Management* 1, no. 1 (2021): 182–192.
- Carretero, Stephanie, Riina Vuorikari, and Yves Punie. *DigComp 2.1. The Digital Competence Framework for Citizens with Eight Proficiency Levels and Examples of Use*, EUR 28558 EN. Luxembourg: Publications Office of the European Union, 2017. [CrossRef]

- Chiu, Shih-Yung, and Hwei-Lin Chuang. "Employability and Wage Compensation in an Asian Economy: Evidence for Female College Graduates in Taiwan." *Emerging Markets Finance and Trade* 52, no. 4 (2016): 853–868.
<https://doi.org/10.1080/1540496X.2015.1117844>
- Edge, Christi, Elizabeth Monske, Stacy Boyer-Davis, Steven VandenAvond, and Brad Hamel. "Leading University Change: A Case Study of Meaning-Making and Implementing Online Learning Quality Standards." *American Journal of Distance Education* 36, no. 1 (2022): 53–69.
<https://doi.org/10.1080/08923647.2021.2005414>
- European Commission, Directorate-General for Education, Youth, Sport and Culture, Stine Wollscheid, Bjørn Stensaker, and Ben Jongbloed. *Dropout and Completion in Higher Education in Europe: Main Report*. Luxembourg: Publications Office of the European Union, 2015.
<https://data.europa.eu/doi/10.2766/826962>
- European Commission, EACEA, and Eurydice. *Modernisation of Higher Education in Europe: Access, Retention and Employability*. Luxembourg: Publications Office of the European Union, 2014.
- European Commission, EACEA, and Eurydice. *The European Higher Education Area in 2018: Bologna Process Implementation Report*. Luxembourg: Publications Office of the European Union, 2018.
- European Skills, Competences, Qualifications and Occupations (ESCO). *Az egész életen át tartó tanúláshoz szükséges kulcskompetenciák: Európai referenciakeret*. 2009.
<http://ofi.hu/tudastar/nemzetkozi-kitekintes/egesz-eleten-at-tarto>
- Dringó-Horváth, Ida, Judit Dombi, László Hülber, Zsófia Menyhei, Tibor M. Pintér, and Adrienn Papp-Danka. *Az oktatásinformatika módszertana a felsőoktatásban*. Budapest: Károli Gáspár Református Egyetem IKT Kutatóközpontja, 2020.
- Elliott, Ian C., Karin A. Bottom, and Karl O'Connor. "The Status of Public Administration Teaching in the UK." *Journal of Public Affairs Education* 29, no. 3 (2023): 262–274.
<https://doi.org/10.1080/15236803.2023.2202609>
- Farid, Shahid, Rodina Ahmad, and Mujahid Alam. "A Hierarchical Model for E-learning Implementation Challenges using AHP." *Malaysian Journal of Computer Science* 28, no. 3 (2015): 166–188.
- Ferrari, Anusca. *DigComp: A Framework for Developing and Understanding. Digital Competence in Europe*. Brüssel: European Commission, 2013.
- Ferri, Fernando, Patrizia Grifoni, and Tiziana Guzzo. "Online Learning and Emergency Remote Teaching: Opportunities and Challenges in Emergency Situations." *Societies* 10, no. 4 (2020): 86.
<https://doi.org/10.3390/soc10040086>
- Fryer, Thomas, Gary R. McKinney, and Joseph E. Trimble. *Tools and Techniques for Program Improvement: A Handbook for Program Review and Assessment of Student Learning*. Bellingham, WA: Western Washington University, 2006.
- Jenkins, Henry, Sam Ford, and Joshua Green. *Spreadable Media: Creating Value and Meaning in a Networked Culture*. New York: New York University Press, 2013.
- Kerby, Molly B. "Toward a New Predictive Model of Student Retention in Higher Education." *Journal of College Student Retention: Research, Theory & Practice* 17, no. 2 (2015): 138–161.
<https://doi.org/10.1177/1521025115578229>

-
- Monske, Elizabeth, Stacy Boyer-Davis, Steven VandenAvond, and Brad Hamel. "Leading University Change: A Case Study of Meaning-Making and Implementing Online Learning Quality Standards." *American Journal of Distance Education* 36, no. 1 (2022): 53–69.
<https://doi.org/10.1080/08923647.2021.2005414>
- Nemeskeri, Zsolt, Janos Szello, Ivan Zadori, and Eszter Barakonyi. "Digital Competencies and Career Orientation in the 21 st Century: Hungarian Labor Market Approach." *Revista Romana De Statistica* 60, no. 12 (2016): 78–92.
- Olivier, Denise, Brendon Freeman, Christopher Young, Shuyan Yu, and Gaurav Verma. *Employer Satisfaction Survey: Final Report. Report for the Department of Education*. Sydney: University of Sydney Business School, 2014.
https://www.qilt.edu.au/docs/default-source/default-document-library/ess_final_report_june_14_0.pdf
- O'Leary, Rosemary. "Advice to New Teachers: Turn It Inside Out." *PS: Political Science & Politics* 35, no. 1 (2002): 91–92.
<https://doi.org/10.1017/S1049096502000215>
- O'Neill, Michael A. "MPA/MPP Programmes and the Future of Work: Does MPA Curriculum Align with Civil Service Requirements." *Teaching Public Administration* 40, no. 3 (2021): 014473942098450.
<https://doi.org/10.1177/0144739420984503>
- M. Pintér, Tibor. "Digitális kompetencia a felsőoktatásban." [Digital competence in higher education.] *Modern Nyelvoktatás* 25, no. 1 (2019): 47–58.
- Magyarország Kormánya. *Magyarország Digitális Oktatási Stratégiája*. 2016.
<https://digitalisjoletprogram.hu/hu/tartalom/dos-magyarorszag-digitalis-oktatasi-strategiaja>
- Partnership for 21st Century Learning (P21). *Framework for 21st Century Learning*. 2014.
<https://www.battelleforkids.org/networks/p21/frameworks-resources>
- Prensky, Marc. "Digital Natives, Digital Immigrants Part 1." *On the Horizon* 9, no. 5 (2003): 1–6.
<https://doi.org/10.1108/10748120110424816>
- Ramírez-Hurtado, José M., Alfredo G. Hernández-Díaz, Ana D. López-Sánchez, and Víctor E. Pérez-León. "Measuring Online Teaching Service Quality in Higher Education in the COVID-19 Environment." *International Journal of Environmental Research and Public Health*, 18, no. 5 (2021): 2403.
<https://doi.org/10.3390/ijerph18052403>
- Rogers, Donna L. "A Paradigm Shift: Technology Integration for Higher Education in the New Millennium." *Educational Technology Review. AACE Review (formerly AACE Journal)* 1, no. 13 (2000): 19–33.
- Rubaii, Nadia. "Bringing the 21st-Century Governance Paradigm to Public Affairs Education: Reimagining How We Teach What We Teach." *Journal of Public Affairs Education* 22, no. 4 (2016): 467–482.
<https://doi.org/10.1080/15236803.2016.12002261>
- Santos, Helena, João Batista, and Rui Pedro Marques. "Digital transformation in higher education: the use of communication technologies by students." *Procedia Computer Science* 164, (2019): 123–130.
<https://doi.org/10.1016/j.procs.2019.12.163>

- Sarda, Elisa, Olga Kasatkina, and Erica de Vries. "How do Lecturers Conceptualise Pedagogical Innovations in Higher Education?" *Innovations in Education and Teaching International* 61, no. 4 (2023): 611–621.
<https://doi.org/10.1080/14703297.2023.2205871>
- Spady, William G. "Dropouts from Higher Education: Toward an Empirical Model." *Interchange* 2, no. 3 (1971): 38–62.
- Schwab, Klaus. *The Fourth Industrial Revolution*. New York: Crown Publ. Group, 2017.
- Szabó, Szilvia. *Kompetencia alapú emberi erőforrás gazdálkodás* [Competency-based human resource management]. Budapest: Nemzeti Közzolgálati Egyetem, 2014.
- Tari, Annamária. *Z generáció*. Budapest: Tercium Könyvkiadó, 2011.
- Tawafak, Ragad M., Awanis B. T. Romli, Ruzaini bin Abdullah Arshah, and Sohail Iqbal Malik. "Framework Design of University Communication Model (UCOM) to Enhance Continuous Intentions in Teaching and E-Learning Process." *Education and Information Technologies* 25, (2020): 817–843.
- Tinto, Vincent. "Dropout from Higher Education: A Theoretical Synthesis of Recent Research." *Review of Educational Research* 45, no. 1 (1975): 89–125.
<https://doi.org/10.2307/1170024>
- Vass, Vilmos. "A tudásgazdaság és a 21. századi kompetenciák összefüggései." *Új Munkaügyi szemle* 1, no. 1 (2020): 30–37.
- Vuorikari, Riina, Yves Punie, Stephanie Carretero Gomez, and Godelieve Van den Brande. *DigComp 2.0: The Digital Competence Framework for Citizens: Update Phase 1, The Conceptual Reference Model*. JRC Science for Policy Report JRC101254. Luxembourg: Publications Office of the European Union, 2016.
<https://doi.org/10.2791/11517>
- Vuorikari, Riina, Stefano Kluzer, Yves Punie. *DigComp 2.2: The Digital Competence Framework for Citizens: With New Examples of Knowledge, Skills and Attitudes*. EUR 31006 EN. Luxembourg: Publications Office of the European Union, 2022.

Appendix

| Area of Competence Indicated by the Framework of Reference | Area of Competence Indicated in the Subject Programme | Note - Implementation on the Course |
|--|--|---|
| 1. Information and Data Management | | |
| 1.1. Browse, search and filter data, information and digital content | Digital research methodology (evaluation and exploitation of information and data repositories, knowledge content) | Familiarisation with databases (international and national), online encyclopaedias, public administration knowledge base. Task 1: Create your own CV, update an existing one. Set your own SMART goals |
| 1.2. Evaluation of data, information and digital content | Digital research methodology (evaluation and exploitation of information and data repositories, knowledge content) | Learning and practising the basics of research methodology, data collection and citation styles. |
| 1.3 Managing data, information and digital content | Digital research methodology (information and data repositories, knowledge assessment and exploitation) | Assignment 2: Write a 2-page text based on the use of a database, referring to data according to the appropriate rules |
| 2. Communication and Cooperation | | |
| 2.1. Interaction through digital technologies | Learning collaborative spaces: rules for internal communication, use of online platforms | Use of university social networking sites (digital course materials library) |
| 2.2. Sharing through digital technologies | Learning collaborative spaces: rules for internal communication, use of online platforms | Task 3: Editing a newsletter in Canva, individually or in a group |
| 2.3. Citizen participation through digital technologies | E-citizenship, e-services, e-government (Client Gateway), e-commerce I-II. | Use of online administrative interfaces, client portal |

| | | |
|--|---|---|
| 2.4. Collaboration through digital technologies | Learning collaborative spaces: rules for internal communication, use of online platforms | Joint editing in Canva (pair and group work) Task 5: Preparation of a test in Kahoot for the lecture material and the relevant textbook chapter (fixed task, week by week) Learning and using an online office suite (Forms, Whiteboard, Sway, One drive) |
| 2.5. General rules of etiquette for network communication (Netiquette) | Digital records, database management Digital space: organisational social media functions (methods, functions, possibilities) | Letter writing, internal communication, netiket |
| 2.6. Digital identity management | E-citizenship, e-services, e-government (Client Gateway), e-commerce I.-II. | Confident use of the client gateway secure use and storage of passwords |
| 3. Digital Content | | |
| 3.1. Creating digital content | Digital literacy I: creating and editing large documents, multimedia and web documents, graphics Content production: Libre Office funds, requests, applications Presentation techniques, methods, programs I. (PPT, Google slides, Prezi, other online interfaces) | Download images (learn about usage rights), image editing, pasting, poster making. Learning about simple image editing programs. Using Libre Office Confident use of Office |
| 3.2. Editing digital content | Digital literacy II: creating and editing large documents, multimedia and web documents, graphics Content production: Libre Office funds, requests, applications Presentation techniques, methods, programs I. (PPT, Google slides, Prezi, other online interfaces) | Download images (learn about usage rights), find images, edit images, paste images, create posters Using Libre Office Confident use of the Office program downloading images (understanding rights of use), editing, pasting, poster making Using Libre Office Confident use of Office Make a presentation |

| | | |
|--|---|--|
| 3.3. Copyright and licences | Digital research methodology: searching for, evaluating and using information and repositories, knowledge content | Review of copyright and licensing, explanation of legislation Problems with the use of AI content |
| 3.4. Programming | | |
| 4. Security | | |
| 4.1. Protection of assets | Use of IT tools, mobile devices, cloud services, utilities | Passwords, settings |
| 4.2. Protection of personal data and privacy | Digital records, database management Security awareness in the digital space | Passwords, settings, backups, cloud services (learning about university cloud services) |
| 4.3. Protecting health and well-being | | |
| 4.4. Environment | | |
| 5. Problem Solving | | |
| 5.1. Solving technical problems | Theory and practice of UX | User needs, online responses |
| 5.2. Formulating needs and technological responses | Theory and practice of UX | Task 5: Website testing and analysis |
| 5.3. Active use of digital technologies | e-Citizenship, e-Services, e-Administration (Client Gateway), e-Commerce I-II. | Digital citizenship, online administration, |
| 5.4 Assessing digital competence gaps, remediation and repetition. | | |

Table 1. Document analysis

| Criterion examined | Spring 2023 | Spring 2024 |
|--------------------|---|--|
| Objective | Measurement of input and output competency levels based on DigCompSat | Evaluation of course effectiveness through document analysis and proprietary questionnaire |
| Method | DigCompSat self-assessment (n=120 input, n=58 output) | Comparison of DigComp 2.1 with the subject program + student questionnaire (n=62) |

| | | |
|----------------------------|---|---|
| Developed competence areas | 5 areas, measured based on DigCompSat | All 5 DigComp areas covered |
| Greatest improvement | Content creation (10.65%), skills (9.43%), security awareness (27.29% among 35–39-year-olds) | Digital content editing (81%), copyright (79%), information filtering (77%) |
| Least progress | Programming, security, problem solving showed a decline among 30–34-year-olds | Programming (29%), environmental protection (32%), health protection (44%) |
| Student satisfaction | Average competency level: 63.3%, positive shift after the course | Average rating: 3.66–4.18 (on a 5-point scale) |
| Demographic effects | No significant differences based on age or prior education | The sample consisted of the same age group, no further details provided |
| Gender differences | Men started at a higher level, women showed greater improvement (average +8.85%) | No details provided |
| Age group effect | 35–39-year-olds improved the most, 30–34-year-olds overestimated themselves | The sample consisted of the same age group, no details provided |
| Conclusion | Competence development is possible in a short period of time, but self-assessment is not reliable | The course fits well with the DigComp framework and is useful for students |

Table 2. Comparison of the results of the two questionnaires (own editing)

Regional and Sectoral Applications of Artificial Intelligence and the Role of Technological Anxiety

This study analyzes the multidisciplinary applications and socio-psychological effects of artificial intelligence (AI). Through a systematic literature review, 150 relevant Q1-rated publications from Scopus and Web of Science were synthesized. Our analysis spans from macroeconomic technological transformations to deep-seated workplace anxiety. Results confirm that while AI primarily helps address shortages of skilled professionals in healthcare and education, adoption is hindered by psychological barriers such as the cognitive replacement crisis and AI-FoMO (fear of technological lag). Regional analysis reveals sharp disparities: Asia leads in sectoral integration and anxiety research, Europe faces a significant digital divide, while infrastructure constraints limit the adoption and scalability of AI in African countries. The study concludes that successful adoption depends not only on technical metrics but also on trust-building, Explainable AI (XAI) transparency, and stable regulatory frameworks.

Keywords: *artificial intelligence, literature review, regional analysis, sectoral applications, technological anxiety*

Acknowledgement

This research was supported by the University Research Scholarship Programme (EKÖP) within the framework of the Cooperative Doctoral Program (EKÖP-KDP-25), funded by the Ministry of Culture and Innovation and the National Research, Development and Innovation Fund.

Author Information

Gergely Lülök (Corresponding Author), PhD Student, Budapest University of Technology and Economics, Department of Management and Business Economics, 2 Magyar tudósok körútja, 1117 Budapest, Hungary

lulok.gergely@edu.bme.hu

Zoltán Sebestyén, Associate Professor, Budapest University of Technology and Economics, Department of Management and Business Economics, 2 Magyar tudósok körútja, 1117 Budapest, Hungary

sebestyen.zoltan@gtk.bme.hu

How to cite this article:

Lülök, Gergely, Zoltán Sebestyén. "Regional and Sectoral Applications of Artificial Intelligence and the Role of Technological Anxiety". *Információs Társadalom* XXV, no. 4 (2025): 94–126.

==== <https://dx.doi.org/10.22503/infars.XXV.2025.4.5> ====

*All materials
published in this journal are licenced
as CC-by-nc-nd 4.0*

1. Introduction

The scientific study of artificial intelligence has grown rapidly in recent years, necessitating the systematic organization of available results. Our current research analyzes 150 Q1-ranked publications selected from the Scopus and Web of Science databases to reveal patterns in the sectoral and regional application of technology. The study is grounded in the observable duality between the operational advantages of the technology (such as increased efficiency or diagnostic accuracy) and the psychological resistance accompanying its application. The aim of our study is to use the examined publications to answer the question of how AI acceptance varies across different segments of the global economy.

In our analysis, we pay particular attention to critical phenomena identified in the sample, such as the need for explainability and specific forms of technological anxiety. In the latter case, based on the articles examined, we distinguish between the fear of cognitive replacement and the phenomenon of AI-FoMO, which appears in recent literature as anxiety about technological backwardness. The structure of the study follows the logic of the extracted data: after establishing the methodological framework, we first present the scientometric characteristics of the publication sources, then provide a detailed synthesis of regional characteristics, with particular attention to the differences between Asian, European, American, and African research focuses. The introduction establishes the foundation for a systematic review that examines the relationship between technical implementation and socio-psychological effects in light of the available empirical data.

2. Methodology

We conducted a multi-stage, systematic literature review following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines to ensure transparency and reproducibility of the analysis. The aim of the study is to comprehensively map the technological, social, and psychological aspects of artificial intelligence. The interrelated steps of data collection, filtering, and categorization are illustrated in Figure 1.

As a first step in data collection, we conducted searches in the international scientific databases Scopus and Web of Science. The research was conducted around two distinct thematic pillars, to which we assigned specific English-language search terms:

- First pillar (regional and sectoral applications): for example, “AI industrial application,” “AI adoption across countries,” “regional AI trends,” “AI in health-care,” “AI in manufacturing,” “AI in finance,” and “industry 4.0 AI.”
- Second pillar (social impacts, anxiety, acceptance, human factors): for example, “AI anxiety,” “fear of artificial intelligence,” “technology acceptance model AI,” “trust in AI,” “algorithmic aversion,” “technostress,” “explainable AI,” and “AI transparency.”

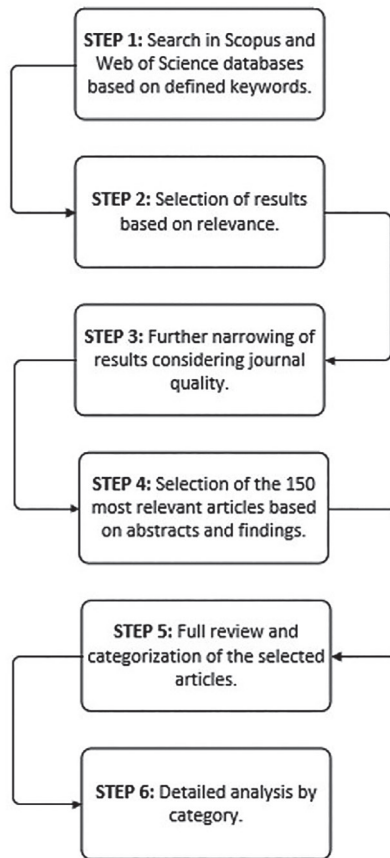


Figure 1. Process of literature selection and analysis (Source: Authors' own elaboration)

The initial search yielded more than 3,000 hits. In the second stage, we significantly narrowed down the number of records through a conscious methodological decision. In doing so, we excluded articles focused purely on mathematical algorithm optimization, engineering code efficiency, or IT architecture.

This was necessary because the focus of the study was not on the internal workings of algorithms or their coding background, but on the practical application of the technology, sectoral wage effects, and human anxiety and acceptance. Since technically focused studies did not provide empirical data on market reactions and the psychology of human-machine interaction, they were not relevant to the present synthesis. To ensure scientific rigor, we limited our selection to journals ranked Q1 according to Scimago Journal Rank (SJR). After a detailed analysis of the abstracts and results, we included 150 publications containing specific empirical data or systematic sectoral analyses in the final synthesis. The selected studies were classified into two main categories: industry and regional trends according to the NACE Rev. 2 classification, and social impacts and technological anxiety.

3. Literature synthesis on the relationship between regional and sectoral adaptation and technological anxiety

The dynamic expansion of the literature on artificial intelligence necessitates the systematization of the results. This study synthesizes the literature from a novel integrative perspective that combines regional, sectoral, and technological anxiety dimensions. In contrast to specific studies focusing on a single sector, our analysis takes an integrated approach. It examines regional and industrial adaptation and the social context of technology simultaneously, with particular attention to the methodological solutions of previous reviews.

3.1. Reviews of the literature to date

The rapid growth in scientific literature on artificial intelligence has given rise to the latest systematic reviews, which have shifted their focus from technological feasibility to organizational integration and trust building (Bolaños et al. 2024). A critical observation regarding *administrative and service support activities* is that developments are often driven by “technological pressure” rather than real needs. As a result, investments often fail to pay off, and successful adaptation requires “decision intelligence” frameworks that optimize human-AI interaction (Di Vaio, Hassan and Alavoine 2022).

In the field of *human health and social work activities*, the technical maturity of diagnostic algorithms is well established, as analyses in the fields of heart disease and oncology have confirmed that the accuracy of these models exceeds that of traditional methods (Ahsan and Siddique 2022; Hoyos, Aguilar and Toro 2021; Nassif et al. 2022). However, these excellent technical metrics cannot be automatically converted into clinical success due to data imbalance, which is why the latest research is already investigating human-machine collaboration models that guarantee patient safety (Andrade et al. 2026). A similar trend can be observed in the *information and communication sector*, particularly in relation to cybersecurity.

The effectiveness of reinforcement learning has been demonstrated in several studies, and the integration of quantum computing and AI promises a breakthrough in drug discovery, although the lack of standardized evaluation frameworks still hinders its widespread adoption (Adawadkar and Kulkarni 2022; Pineda et al. 2026).

In the *education sector*, the literature points out that integration is not an infrastructure issue, but rather a pedagogical and competency issue, where adaptive systems enable personalized education (Chiu et al. 2023; Kabudi, Pappas and Olsen 2021; Lyu et al. 2025). However, the acceptance of technology is hampered by serious psychological barriers and anxiety, which can be addressed by developing “AI literacy” and integrating “prompt engineering” into the curriculum (Bewersdorff et al. 2023; Laupichler et al. 2022; Park 2025). In the field of *financial and insurance activities*, solving the “black box” problem has become a priority; due to the regulatory environment, the use of Explainable AI (XAI) methods is now a legal requirement

to ensure the traceability of decisions (Černevičienė and Kabašinskas 2024; Cil and Yildiz 2025).

In *professional, scientific, and technical activities*, analyses refute fears of full automation, as AI plays a predominantly complementary role, scaling human creativity in research and development and marketing (Bahoo, Cucculelli and Qamar 2023; Johnson et al. 2022; Labib 2024; Truong and Papagiannidis 2022; Gupta et al. 2023; Ha 2022). The fear of identity loss has been identified as a barrier to social acceptance, which can be mitigated by transparent systems and emotional support functions (Jiang et al. 2024; Chaturvedi et al. 2023; Haque, Islam and Mikalef 2023). Finally, in the areas of *transportation and storage*, as well as *electricity, gas, steam and air conditioning supply*, AI-based predictive maintenance is key to safety and sustainability, although scalability remains a challenge at the smart city level (Garcia et al. 2026; Magdin 2025; Rehman 2026; Rashidieranjbar et al. 2026).

3.2. Regional and sectoral comparison

The development and application of artificial intelligence vary around the world, as different continents and countries approach the technology based on their unique social, economic, and technological characteristics. The following summary examines how certain countries and regions relate to AI, what challenges and opportunities they face, and how they integrate it into their own industries and public policies. The analysis aims to highlight the global and local factors that shape AI research and application.

3.2.1. Asia

Research findings from the Asian region highlight that AI has become a key driver of the macroeconomy. Chinese modelling of *professional, scientific and technical activities* has confirmed that the impact of AI on the labour market is not linear. Machine replacement causes short-term wage pressure but generates a significant 15–20% skill premium for highly skilled workers through the emergence of new job roles (Cheng et al. 2024; Qian et al. 2023). Corporate competitiveness is based on technological orientation (Liu et al. 2024), knowledge management fusion, and trust in agile CRM systems (Chatterjee et al. 2021).

In the *financial and insurance activities* sector, LSTM networks provide consistently more accurate forecasts even in volatile market phases (Cheng et al. 2024). Sustainability studies have revealed a peculiar “U-shaped” relationship: AI implementation initially increases environmental impact, but once the efficiency threshold is reached, it becomes an engine of green growth (Zhao, Gao and Sun 2022). In the *manufacturing* industry, AI has brought about a breakthrough in competitiveness in low-tech sectors (Fu et al. 2024; Liu et al. 2020), while in resource-constrained environments, solutions combined with IoT dominate (Qin 2024).

In the *information and communication* sector, awareness of smart technologies directly increases learning motivation (Deng et al. 2025). Within *transportation and warehousing*, deep reinforcement learning has reduced traffic waiting times by 12% (Guo and Wang 2021), while in logistics planning, multi-objective mathematical models minimize environmental impacts (Dini et al. 2025). In *wholesale and retail trade; repair of motor vehicles*, export market forecasts have achieved 98% accuracy (Sohrabpour et al. 2021), and the acceptance of sales systems is most influenced by habits and hedonic motivation (Das and Datta 2024).

Regarding the *electricity, gas, steam and air conditioning supply* sector, hybrid methods help ensure the stability of smart grids (Dai, Yang and Leng 2022). In *water supply; sewerage, waste management and remediation activities*, BiLSTM models have improved the accuracy of flood protection forecasts (Kareem et al. 2022). In *education* in Hong Kong, the main obstacle remains a lack of teacher knowledge (Cheng and Wang 2023), while in the *administrative and support service activities* sector, researchers identified 15 management constraints in the Indian public sector (Misra et al. 2023). In Japan, government cooperation forms the basis of the medical AI strategy (Ishii et al. 2020), although strict budget constraints and the ethical risks of data-driven decision-making pose a global challenge (Zhu, Zhang and Feng 2022; Marda 2018).

The success of technological adaptation is determined by the dynamic balance between collective fear and individual competence. Eight factors of anxiety have been identified in *professional, scientific, and technical activities* (Li and Huang 2020), where avoidance of uncertainty increases fear (Hitsuwari and Takano 2025). The innovative behavior of highly skilled workers is doubly influenced by AI. While technological features stimulate creativity, the “replacement crisis” and technostress directly undermine motivation (Verma and Singh 2022, 2024).

“Fear of falling behind” (AI-FoMO) has emerged, which simultaneously reduces subjective well-being and encourages learning (Yu 2025; Liao, Fang and Li 2026). In the field of *information and communication*, social media increases literacy, but audio content causes information overload and anxiety (Li et al. 2026; Qiu et al. 2025). In *education*, preparedness and anxiety are inversely proportional: awareness of effects increases risk perception (Liang and Zhai 2025), while teacher support relieves tension through technophilia (Chai and Sha 2025).

In language learning, AI assistants reduced inhibitions by 30% (Zhang, Meng and Ma 2024; Wang and Wen 2025), but overuse can lead to academic stress and loneliness (Wang and Xu 2026).

High anxiety among medical students and librarians stems from systemic unpreparedness and a perceived loss of control (Özbek Güven, Yılmaz and Inceoğlu 2024; Gültekin and Kavak 2025). In the *human health and social work activities* sector, the doctor-patient relationship is asymmetrical: patients fear dehumanization, while doctors fear loss of autonomy (Li and Liu 2025; Li et al. 2025; Yang 2026). Anxiety among pediatric nurses is particularly high, fueled by a loss of empathy (Tutar, Özgörü and Ögür 2025; Ünal and Avcı 2024; Burucu and Polat 2025). In contrast, AI support improves the quality of life of cancer patients, although professionals remain wary of machine replacement (Demir-Kaymak et al. 2024; Soysal, Çalışkan and Turgut 2025).

In *financial and insurance activities*, customers' technological anxiety and risk aversion weaken acceptance (Cintamür 2024; Boustani 2021; Ghazwani et al. 2022). In the *manufacturing* industry, the negative effects of technostress can be mitigated by digital accountability and servant leadership (Zhang et al. 2025; Chen et al. 2025; Malik et al. 2021). Within the *accommodation and food service activities* sector, the perception of GenAI can lead to job insecurity and rule-breaking behavior (Zhao, Yuan and Song 2024; Huang et al. 2025). Finally, in *administrative and support service activities*, the anthropomorphic nature of AI and concerns about design capabilities increase dependency and replacement fears (Çetin 2024; Jo and Park 2025; Zhang and Zhang 2025).

3.2.2. Europe

The most striking finding of European research is the deep territorial divide. Due to the gap between the “expert” northwestern and “novice” eastern regions, the EU cannot be considered a unified digital space. In the field of *professional, scientific and technical activities*, AI-based models in advanced ecosystems already predict R&D performance with high accuracy (de la Paz-Marín, Campoy-Muñoz and Hervás-Martínez 2012), and industrial AI (IAI) has been shown to catalyze innovation in the ten leading European economies (Anser and Adebayo 2026). The development of startup ecosystems is determined by the level of adaptation of neighboring regions and integration into local technology clusters (Vicente, Álvarez-Rodríguez and Suárez-Álvarez 2025; Vásquez, Sandulli and Gallego 2026).

In the *manufacturing* industry, AI supports the transition to a circular economy by increasing resource efficiency (Sjödín, Parida and Kohtamäki 2023), although the innovation gap remains drastic. While 24% of companies in Denmark use the technology, only 1.4% in Romania do so (Brodny and Tutak 2021, 2022, 2024). Energy security remains the top priority in the areas of *electricity, gas, steam and air conditioning supply*, where federated platforms help test reliable algorithms (Liu et al. 2025; Sarmas et al. 2026).

In the *financial and insurance activities* sector, CatBoost models more effectively predict corporate bankruptcy, while in green finance, algorithms optimize sustainable investments (Jabeur et al. 2021; Wang et al. 2025).

In the *water supply; sewerage, waste management and remediation activities* sector, deep learning aids climate research by detecting Arctic thermokarst formations (Vasile et al. 2025), while in French hospitals, green digital learning improves environmental performance (Benzidia, Makaoui and Bentahar 2021). In *administrative and support service activities* and HR, based on Scandinavian examples, gamification and automated recruitment through chatbots reduce bias (Allal-Chérif, Yela Aránega and Castaño Sánchez 2021; Allal-Chérif 2022; Trocin et al. 2021). In the SME sector, the key to success is the development of internal digital competencies (Arranz et al. 2023). In *human health and social work activities*, EU sustainability is underpinned by federated learning roadmaps (Kommusaar et al. 2026).

In this region, acceptance is not merely a technical process, but a deeply rooted social one, in which trust plays a decisive mediating role. In *professional, scientific and technical activities*, technological self-efficacy directly reduces fear: a higher sense of competence increases trust in automated systems (Montag et al. 2023). AI literacy increases the perception of usefulness while reducing socio-cultural concerns (Schiavo, Businaro and Zancanaro 2024). At the same time, in the field of *administrative and support service activities*, the digital divide translates into technological fear: in eastern regions, 70–90% of the population fears the automated replacement of jobs (Vasilescu et al. 2020).

Among HR professionals, there is tension between the threat and the opportunity for skill enhancement (Guggemos 2024), while the phenomenon of AI-FoMO in the workplace is generated by the loss of decision-making autonomy and the fear of skills becoming obsolete (Méndez-Suárez, Ćukušić and Ninčević-Pašalić 2026). In the *information and communication* sector, a study of the Czech public confirmed that age and education significantly influence the ability to recognize AI content (Moravec et al. 2024). In marketing management, the exploitation of technology is hampered by ethical concerns among managers (Volkmar, Fischer and Reinecke 2022; Capatina et al. 2020), while in the *manufacturing* industry, employee resistance can prevent data-driven insights from being converted into practical sustainability measures (Garcés-Marín, Arias-Pérez and Restrepo-Estrada 2026). Finally, social attitudes are shaped by fundamental values such as altruism and the desire for equality (Kanzola, Papaioannou and Petrakis 2024), while managers remain responsible for green talent management and minimizing uncertainty (Odugbesan et al. 2022).

3.2.3. United States and Australia

The research focus in the United States and Australia is based on the tension between high-level technological developments and social-regulatory constraints. In the field of *human health and social work activities*, deep learning models have already achieved human-level accuracy in medical imaging and diagnostics (Esteva et al. 2019).

In Australia, AI-based triage systems show 98.3% agreement with specialist decisions (Nguyen et al. 2025), and screening for diabetic retinopathy not only improves quality of life but also results in significant cost savings (Hu et al. 2024). Within the Australian Epilepsy Project, a responsible AI framework based on multimodal data supports precision diagnostics (Pedersen et al. 2025).

At the same time, regulation in the US is slow: by 2019, only two complex algorithms had been approved as medical devices (Cheung et al. 2019). However, the corrective potential of AI can mitigate racial disparities in care: an American model has confirmed that algorithms measure patients' pain levels more objectively than doctors, regardless of their skin color (The Lancet Digital Health 2021). In public health, analyzing social media data is effective in predicting the spread of sexually transmitted infections (Young, Crowley and Vermund 2021).

In *professional, scientific and technical activities* in the US, XAI provides more accurate insight into supply chain volatility (Berger 2023). However, in *financial and insurance activities*, operational incidents are severe: technological failures cause an average cumulative yield loss of 21.04% and an increased risk of bankruptcy for service providers (Durongkadej, Hu and Wang 2024). In the *electricity, gas, steam and air conditioning supply* sector, an XAI analysis of Australian cities revealed that the suitability of solar panel installation is more strongly influenced by population distribution and forest cover than by the amount of solar radiation (Shen et al. 2026). In the case of *administrative and support service activities*, an “application paradox” can be observed: despite technical installation, actual use is inhibited by fear of failure (Sarkar 2025).

Technological adaptation is accompanied by deep psychological and political divisions in both regions. In the Australian auditing sector, the “black box phenomenon” and the opacity of algorithms are causing mistrust in *professional, scientific and technical activities* (Yang, Blount and Amrollahi 2024). In the US labor market, the AI shock has triggered political radicalization: groups threatened by technologization are responding to automation with conservative social and left-wing economic views (Jacobs 2024). Psychological resistance is fueled by a prevention-focused mindset, which reinforces fears of job replacement (Sapru 2026).

In the *accommodation and food service activities* sector, anxiety about the introduction of algorithms significantly increases the intention of US hotel employees to leave their jobs (Koo, Curtis and Ryan 2021). In the *human health and social work activities* sector in Australia, the acceptance of AI-based mobile applications is asymmetrical: users are more willing to share data on heart disease than information related to mental health (Vo et al. 2025). In addition, an examination of generative AI revealed deep-rooted biases: 90.8% of images depicting paramedics were male and 90.5% were Caucasian, which differs significantly from the actual, diverse workforce composition (Currie, Hewis and Ebbs 2025). In the field of *information and communication*, the fear of being left out of AI (AI-FoMO) has been validated in the US, which particularly affects the well-being of women and young people, while technological literacy has proven to be an effective protective factor against it (Yu 2025).

3.2.4. Africa and global comparative trends

In Africa, algorithms have become critical tools for bridging the gap in *human health and social work activities*. The 92.25% sensitivity achieved in screening for diabetic retinopathy makes it possible to radically reduce the risk of vision loss even without the presence of a specialist (Mathenge 2019). Mobile-based digital systems have been shown to improve HIV prevention reach indicators and the effectiveness of clinical toxicology diagnoses in Ghana, Kenya, and Nigeria (Mbewe et al. 2026; Teferi et al. 2025). In oncology care, AI offers paradigm-shifting opportunities, but its diagnostic potential is hampered by severe infrastructure constraints and a lack of local data assets (Dako et al. 2025; Hughes et al. 2026; Standley et al. 2025).

In the *manufacturing* industry, negative labor market correlations are evident: automation shock accounts for nearly 30% of employment variance, resulting in short-term job losses (Giwa and Ho 2026). In the South African public sector, the lack of specific AI skills in *administrative and support service activities* hinders efficiency gains (Sikhakhane 2025). Meanwhile, in the *information and communication* sector, professional skepticism is managed by harmonizing technology with local value systems, such as the Ubuntu philosophy (Gondwe 2024).

At the global level, the dynamics of *professional, scientific and technical activities* are shaped by differences in research volume and regulatory maturity. While the US produces orders of magnitude more publications on the use of electronic health records, Chinese networks show more intensive international cooperation (Chen et al. 2018). In terms of medical device approval, the US and Europe have adopted algorithms at a similar pace, but the lack of specific regulatory pathways undermines transparency (Muehlematter, Daniore and Vokinger 2021). In the *information and communication* sector, the cost-cutting role of AI dominates (Balmer, Levin and Schmidt 2020). In predicting the attractiveness of the business environment, Cat-Boost models have confirmed that a stable government background is the primary draw for technological investments (Jabeur et al. 2022).

In the *manufacturing* industry, a fusion of Industry 4.0 and sustainability is observed: Asia focuses on robotics, while Europe focuses more on direct environmental goals (Calabrese et al. 2023). The global transformation of *education* systems is moving toward personalization (Bhutoria 2022). In the areas of *electricity, gas, steam and air conditioning supply*, the performance of the renewable sector is asymmetrically dependent on technological shocks (Urom et al. 2022).

During the analysis of the research corpus, a striking regional gap was identified: none of the studies in the sample dealt with the issue of technological anxiety in Africa. This suggests that research on the continent currently focuses primarily on technology accessibility and overcoming implementation barriers, pushing the psychological consequences into the background. In contrast, global patterns highlight that in the *accommodation and food service activities* sector, revealing AI's presence in hedonistic processes can cause a loss of trust due to an existential threat (Xu et al. 2024).

In *wholesale and retail trade; repair of motor vehicles*, consumer preferences are segmented; the human-like design of AI products in Japan and the US reinforces data privacy concerns (Frank 2024). Finally, meta-analyses in *human health and social work activities* confirm the supportive power of the technology: AI assistants are significantly capable of reducing the pain and anxiety of cancer patients (Sezgin and Bektas 2025).

4. Quantitative analysis of the examined publication stock

4.1. Changes in research activity over time

An analysis of the publication dates of the 150 studies included in the review reveals the dynamics of scientific discourse. Based on the data, research intensity has become dominant in the last five years following a longer preparatory phase (Table 1).

| Year of publication | Quantity |
|---------------------|----------|
| 2012 | 1 |
| 2018 | 2 |
| 2019 | 3 |
| 2020 | 6 |
| 2021 | 16 |
| 2022 | 21 |
| 2023 | 14 |
| 2024 | 31 |
| 2025 | 36 |
| 2026 | 20 |

Table 1. Yearly distribution of analyzed publications (2012–2026)
(Source: own elaboration)

Between 2012 and 2018, publication activity was still sporadic; only one or two relevant studies per year could be identified in the sample examined. At this stage, the application of the technology was still limited to a narrow circle, and research focused primarily on basic feasibility issues.

The turning point occurred in 2021, when the number of publications began to rise dramatically (16), indicating the wider spread of the technology in the industry. Activity peaked in 2024 (31 publications) and 2025 (36 publications), confirming that attention has shifted from technical details to practical issues. The 20 publications already recorded for 2026 are particularly significant: these data suggests that by the end of the year, scientific output is expected to exceed all previous levels, reflecting the critical importance of AI across sectors.

4.2. Scientometric analysis of publication sources

We examined the recognition of scientific journals included in the research along two dimensions: the Hirsch index, which indicates global prestige, and the total number of citations, which indicates the actual impact within the sample. Table 2 shows the ranking of the 10 journals with the highest H-index.

The data show that artificial intelligence research has been published in the world's most prestigious medical and scientific journals. The list is led by Nature Medicine with an outstanding value of 653, followed by the multidisciplinary PLoS ONE (467) and the oncology-focused The Lancet Oncology (432). The presence of Science of The Total Environment (399) in fourth place in the ranking indicates the relevance of the technology to environmental science. The social sciences have strengthened their position in the middle of the list: in addition to the new entry

Social Science & Medicine (296), the high index of the Journal of Business Research (292) and the psychology-focused Computers in Human Behavior (275) demonstrates the importance of research into human-machine interaction and social impacts. The bottom of the ranking is closed by more specific technological applications (Renewable Energy – 270, Neurocomputing – 216) and Technological Forecasting and Social Change (209), which forms the main basis of the present research.

| Journal | H-index |
|---|---------|
| Nature Medicine | 653 |
| PLoS ONE | 467 |
| The Lancet Oncology | 432 |
| Science of The Total Environment | 399 |
| Social Science & Medicine | 296 |
| Journal of Business Research | 292 |
| Computers in Human Behavior | 275 |
| Renewable Energy | 270 |
| Neurocomputing | 216 |
| Technological Forecasting and Social Change | 209 |

Table 2. Ranking of source journals by H-index (Source: own elaboration)

However, the actual influence within the research field is more accurately reflected by the total number of citations, as shown in Table 3. This comparison reveals a significant shift in the global H-index ranking.

| Journal | Total citations |
|--|-----------------|
| Technological Forecasting and Social Change | 7546 |
| Nature Medicine | 6026 |
| Computers and Education: Artificial Intelligence | 4900 |
| California Management Review | 4756 |
| International Journal of Information Management | 3995 |
| Journal of Industrial Information Integration | 1803 |
| Journal of Management Analytics | 1010 |

| | |
|-------------------------------------|-----|
| Artificial Intelligence in Medicine | 930 |
| The Lancet Digital Health | 774 |
| Technology in Society | 719 |

Table 3. Ranking of source journals by total citations within the sample
(Source: own elaboration)

Technological Forecasting and Social Change occupies the leading position: the journal not only has the largest publication volume (42) but also dominates in terms of total citations with 7,546 references. Nature Medicine retained its second place with a single publication (6,026 citations), which indicates the outstanding significance of the medical results published. Computers and Education: Artificial Intelligence ranks third with 4,900 citations, reflecting the explosive interest in educational technology. This is followed by leading management and information systems journals: the high citation rates of California Management Review (4,756) and the International Journal of Information Management (3,995) show that corporate decision-making and information management are among the most active areas of research. The second half of the list includes journals on industrial digitization (Journal of Industrial Information Integration) and business analytics (Journal of Management Analytics). The ranking is rounded off by three journals with a more specific focus: Artificial Intelligence in Medicine and The Lancet Digital Health in the field of medical informatics, and Technology in Society, which examines the social embeddedness of technology.

4.3. Regional and sectoral distribution

The third pillar of bibliometric analysis is the examination of the geographical origin of publications and the economic sectors involved. A cross-sectional analysis of the data reveals that the adaptation of artificial intelligence is not proceeding evenly across the global economy, but rather that different regions have different areas of focus and specialization (Table 4).

| Region | NACE Rev. 2 sector | Percentage / Sector | Percentage / Region |
|--------|---|---------------------|---------------------|
| Africa | Administrative and support service activities | 0.8% | 7.3% |
| | Manufacturing | 0.8% | |
| | Human health and social work activities | 4.9% | |
| | Information and communication | 0.8% | |

| | | | |
|---|--|-------|-------|
| Australia | Human health and social work activities | 4.1% | 5.7% |
| | Professional, scientific and technical activities | 0.8% | |
| | Electricity, gas, steam and air conditioning supply | 0.8% | |
| Asia | Administrative and support service activities | 1.6% | 43.9% |
| | Manufacturing | 4.1% | |
| | Human health and social work activities | 4.9% | |
| | Information and communication | 2.4% | |
| | Wholesale and retail trade; repair of motor vehicles | 2.4% | |
| | Education | 5.7% | |
| | Financial and insurance activities | 1.6% | |
| | Professional, scientific and technical activities | 16.3% | |
| | Accommodation and food service activities | 1.6% | |
| | Transportation and storage | 1.6% | |
| | Electricity, gas, steam and air conditioning supply | 0.8% | |
| Water supply; sewerage, waste management and remediation activities | 0.8% | | |
| Europe | Administrative and support service activities | 2.4% | 26.0% |
| | Manufacturing | 2.4% | |
| | Human health and social work activities | 0.8% | |
| | Financial and insurance activities | 0.8% | |
| | Professional, scientific and technical activities | 17.9% | |
| | Electricity, gas, steam and air conditioning supply | 1.6% | |

| | | | |
|-------------------|--|------|------|
| USA | Administrative and support service activities | 0.8% | 8.9% |
| | Human health and social work activities | 0.8% | |
| | Information and communication | 0.8% | |
| | Financial and insurance activities | 4.1% | |
| | Professional, scientific and technical activities | 0.8% | |
| Cross-continental | Manufacturing | 0.8% | 8.1% |
| | Human health and social work activities | 0.8% | |
| | Wholesale and retail trade; repair of motor vehicles | 0.8% | |
| | Education | 0.8% | |
| | Professional, scientific and technical activities | 3.3% | |
| | Accommodation and food service activities | 0.8% | |
| | Electricity, gas, steam and air conditioning supply | 0.8% | |

Table 4. Regional and sectoral distribution of research activity
(Source: own elaboration)

Looking at geographical distribution, Asia clearly dominates, accounting for 43.9% of the total sample. This region also has the most diverse sectoral portfolio. The backbone of research here is professional, scientific, and technical activities, but there is also significant activity in education, human health care, and manufacturing. This diversified presence suggests that technology in the Asian region has already moved beyond the experimental phase and is undergoing widespread socio-economic integration.

Europe is the second largest research center with a 26% share. However, the continent's profile is significantly more concentrated than Asia's: most publications fall into the category of professional, scientific, and technical activities. This pattern indicates that the emphasis in Europe continues to be strongly on research and development and theoretical foundations, while practical industrial applications are less represented in the sample examined.

Africa and Australia exhibit a particularly interesting pattern. Although their global share is lower (7.3% and 5.7%, respectively), their research profile differs markedly from the global average. In both regions, human health and social care is the dominant sector. In Africa, most of the publications examined focus on

healthcare applications, as do a significant proportion of those in Australia. This suggests that in these regions, the primary drivers of artificial intelligence are social welfare and improving the efficiency of the healthcare system.

In the case of the United States, the distribution of publications, which account for 8.9% of the sample, is more balanced, with healthcare and information and communication appearing alongside professional activities. Studies resulting from global cooperation spanning multiple continents typically examine general scientific and technical issues as well as global challenges (e.g., climate change, energy).

Studies resulting from global cooperation spanning multiple continents account for 8.1% of the total sample.

The sectoral analysis according to NACE Rev. 2 highlights the focus areas of research related to artificial intelligence. Based on the data, more than 40% of the publications examined can be linked to a single dominant sector (Table 5).

| NACE Rev. 2 sector | Quantity | Percentage |
|--|----------|------------|
| Administrative and support service activities | 8 | 5.3% |
| Manufacturing | 10 | 6.7% |
| Human health and social work activities | 25 | 16.7% |
| Information and communication | 7 | 4.7% |
| Wholesale and retail trade; repair of motor vehicles and motorcycles | 4 | 2.7% |
| Education | 14 | 9.3% |
| Financial and insurance activities | 6 | 4.0% |
| Professional, scientific and technical activities | 62 | 41.3% |
| Accommodation and food service activities | 3 | 2.0% |
| Transportation and storage | 4 | 2.7% |
| Electricity, gas, steam and air conditioning supply | 6 | 4.0% |
| Water supply; sewerage, waste management and remediation activities | 1 | 0.7% |

Table 5. Sectoral distribution of publications by NACE Rev. 2
(Source: own elaboration)

Based on the data on sectoral distribution, the primary basis for research activity is professional, scientific, and technical activities, which, with a share of 41.3%, stand out from the field, confirming that the topic of AI anxiety and acceptance is most prevalent at the level of R&D projects and theoretical methodologies. The second most significant area is human healthcare (16.7%), which reflects critical points of contact in medical diagnostics and patient care, while the education sector (9.3%) indicates the academic spread of technology and the importance of digital literacy.

The fourth strongest pillar is manufacturing (6.7%), where the workplace impacts of automation and robotization generate scientific discourse.

In contrast, the volume of research in the other sectors examined is particularly low. This fragmented distribution highlights the fact that while intensive analysis is taking place in the technology and human-centered basic sectors, in-depth research into the psychological effects in the service industry and public utilities is still in its infancy.

4.4. Patterns of AI anxiety and acceptance

Based on sectoral distribution, uncertainty is most prevalent where human expertise is directly affected by automation. Table 6 below details these correlations:

| Region | NACE Rev. 2 sector | Percentage / Sector | Percentage / Region |
|-----------|--|---------------------|---------------------|
| Australia | Professional, scientific and technical activities | 2.5% | 2.5% |
| Asia | Administrative and support service activities | 2.5% | 67.5% |
| | Manufacturing | 5.0% | |
| | Human health and social work activities | 12.5% | |
| | Information and communication | 2.5% | |
| | Wholesale and retail trade; repair of motor vehicles | 2.5% | |
| | Education | 10.0% | |
| | Financial and insurance activities | 5.0% | |
| | Professional, scientific and technical activities | 25.0% | |
| | Accommodation and food service activities | 2.5% | |
| Europe | Professional, scientific and technical activities | 15.0% | 15.0% |
| USA | Information and communication | 2.5% | 7.5% |
| | Professional, scientific and technical activities | 2.5% | |
| | Accommodation and food service activities | 2.5% | |

| | | | |
|-------------------|--|------|------|
| Cross-continental | Human health and social work activities | 2.5% | 7.5% |
| | Wholesale and retail trade; repair of motor vehicles | 2.5% | |
| | Accommodation and food service activities | 2.5% | |

Table 6. Patterns of AI anxiety and acceptance (Source: own elaboration)

According to the research results, the study of technological anxiety continues to be overwhelmingly prevalent in the Asian region, where 67.5% of relevant publications originated. This dominance is closely related to the region's extremely intense technological expansion and the economic environment in which the use of AI has become a fundamental expectation at all levels of the service sector. Europe's 15% share indicates a stable research presence, while the US (7.5%) and multi-continental analyses (7.5%) indicate a diversification of global discourse.

An analysis of the data in the table shows that anxiety in the field of professional, scientific, and technical activities is the most significant, as this category accounts for nearly 40% of the sample globally. This confirms that the scientific community's focus is on the resistance of highly skilled intellectual workers to AI. Particularly noteworthy is the strengthening of the healthcare and education sectors, reflecting the fact that in professions where human interaction and trust are critical factors, issues of AI acceptance are a priority.

5. Conclusions and summary findings

The temporal evolution of research activity highlights a critical paradigm shift: the explosive, exponential growth following 2019—which peaked in 2024—indicates that AI research has synchronized with real-time industrial implementation. The focus of scientific discourse has shifted from proving technical feasibility to organizational integration and trust building. This is also supported by the prestige of the sources: while medical maturity is confirmed by the presence of *Nature Medicine*, *Technological Forecasting and Social Change* has become the most influential forum for the analysis of strategic and social impacts.

An analysis of geographical distribution shows clear Asian dominance. This dominance is not only quantitative but also qualitative: Asia has the most diversified sectoral portfolio, with AI being actively explored in almost all economic sectors, indicating the deep socio-economic integration of the technology. In contrast, Europe's profile is more concentrated, with research continuing to focus on professional, scientific, and technical activities, i.e., theoretical and methodological foundations.

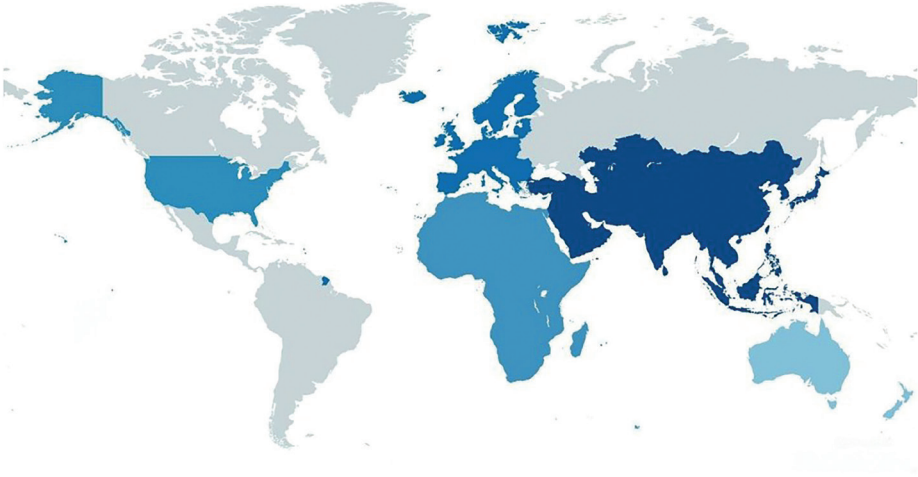


Figure 2. Global research intensity by region
(Source: own elaboration based on MapChart)

The differences in color intensity shown in Figure 2 directly reflect the geographical distribution of the publication stock examined: darker tones indicate the highest research activity, while the number of studies analyzed gradually decreases towards lighter shades. This visual hierarchy highlights the shift in technological focus; the deepest blue color in the Asian region clearly indicates the shift in global research focus to the East, where 43.9% of the sample is concentrated.

Based on the sample studied, the phenomenon of “technological leap” can be observed in developing regions:

In Africa and Australia, research is driven by social constraints, where the primary function of AI is to bridge the shortage of healthcare professionals and increase diagnostic efficiency.

The research focus in the United States is on exploring the tension between innovation breakthroughs and regulatory burdens.

One of the key findings of the research refutes the assumption that fear of AI is limited to less educated groups. According to the data, issues of anxiety and acceptance are most prominent in professional, scientific, and technical occupations. This suggests that the “cognitive replacement crisis” is most intensely felt by the highly educated workforce.

Based on sectoral data, a new correlation can be identified: there is a direct proportionality between research intensity and employees’ sense of threat in intellectual occupations. During the synthesis, the following key correlations emerged:

Interpersonal barrier: In healthcare and education, the primary source of anxiety is not technical inaccuracy, but fear of dehumanization. The solution here is the “augmented intelligence” model, where AI does not replace but supports human decision-making.

Asian testing ground: The high number of anxiety studies in Asia indicates that the fastest technological integration triggers the most intense scientific reflection and the need to build trust.

Need for transparency: In the financial and professional sectors, resolving the “black box” phenomenon, i.e., the use of explainable AI, is no longer an option but a basic condition for legal and social acceptance.

Overall, the publications examined demonstrate that the future of artificial intelligence depends not on further refinement of technical metrics, but on stabilizing the ethical and psychological framework of human-machine symbiosis.

References

- Adawadkar, Amrin Maria Khan, and Nilima Kulkarni. “Cyber-Security and Reinforcement Learning — A Brief Survey.” *Engineering Applications of Artificial Intelligence* 114 (2022): 105116.
<https://doi.org/10.1016/j.engappai.2022.105116>
- Ahsan, Md. Milon, and Zahed Siddique. “Machine Learning-based Heart Disease Diagnosis: A Systematic Literature Review.” *Artificial Intelligence in Medicine* 128 (2022): 102289.
<https://doi.org/10.1016/j.artmed.2022.102289>
- Allal-Chérif, Oihab, Ana Yela Aránega, and Rafael Castaño Sánchez. “Intelligent Recruitment: How to Identify, Select, and Retain Talents from Around the World Using Artificial Intelligence.” *Technological Forecasting and Social Change* 169 (2021): 120822.
<https://doi.org/10.1016/j.techfore.2021.120822>
- Allal-Chérif, Oihab. “Intelligent Cathedrals: Using Augmented Reality, Virtual Reality, and Artificial Intelligence to Provide an Intense Cultural, Historical, and Religious Visitor Experience.” *Technological Forecasting and Social Change* 178 (2022): 121604.
<https://doi.org/10.1016/j.techfore.2022.121604>
- Andrade, João Brainer Clares, Thiago S. Carneiro, George N. Nunes Mendes, Joao Pedro Nardari dos Santos, and Jussie Correia Lima. “A Clinical-AI Correlation for Integrating Artificial Intelligence into Stroke Care: A Systematized Literature Review and Practice Framework.” *International Journal of Medical Informatics* 208 (2026): 106233.
<https://doi.org/10.1016/j.ijmedinf.2025.106233>
- Anser, Muhammad Khalid, and Tomiwa Sunday Adebayo. “AI-driven Renaissance: Revolutionizing Technology Through Industrial Artificial Intelligence in European Economies.” *Technology in Society* 84 (2026): 103129.
<https://doi.org/10.1016/j.techsoc.2025.103129>
- Arranz, Carlos F. A., Marta F. Arroyabe, Nieves Arranz, and Juan Carlos Fernandez de Arroyabe. “Digitalisation Dynamics in SMEs: An Approach from Systems Dynamics and Artificial Intelligence.” *Technological Forecasting and Social Change* 196 (2023): 122880.
<https://doi.org/10.1016/j.techfore.2023.122880>
- Bahoo, Salman, Marco Cucculelli, and Dawood Qamar. “Artificial Intelligence and Corporate Innovation: A Review and Research Agenda.” *Technological Forecasting and Social Change* 188 (2023): 122264.
<https://doi.org/10.1016/j.techfore.2022.122264>

-
- Balmer, Roberto E., Stanford L. Levin, and Stephen Schmidt. "Artificial Intelligence Applications in Telecommunications and Other Network Industries." *Telecommunications Policy* 44, no. 6 (2020): 101977.
<https://doi.org/10.1016/j.telpol.2020.101977>
- Benzidia, Smail, Naouel Makaoui, and Omar Bentahar. "The Impact of Big Data Analytics and Artificial Intelligence on Green Supply Chain Process Integration and Hospital Environmental Performance." *Technological Forecasting and Social Change* 165 (2021): 120557.
<https://doi.org/10.1016/j.techfore.2020.120557>
- Berger, Theo. "Explainable Artificial Intelligence and Economic Panel Data: A Study on Volatility Spillover Along the Supply Chains." *Finance Research Letters* 54 (2023): 103757.
<https://doi.org/10.1016/j.frl.2023.103757>
- Bewersdorff, Arne, Xiaoming Zhai, Jessica Roberts, and Claudia Nerdel. "Myths, Mis- and Preconceptions of Artificial Intelligence: A Review of the Literature." *Computers and Education: Artificial Intelligence* 4 (2023): 100143.
<https://doi.org/10.1016/j.caeai.2023.100143>
- Bhutoria, Aditi. "Personalized Education and Artificial Intelligence in the United States, China, and India: A Systematic Review Using a Human-In-The-Loop Model." *Computers and Education: Artificial Intelligence* 3 (2022): 100068.
<https://doi.org/10.1016/j.caeai.2022.100068>
- Bolaños, Francisco, Angelo Salatino, Francesco Osborne, and Enrico Motta. "Artificial Intelligence for Literature Reviews: Opportunities and Challenges." *Artificial Intelligence Review* 57, no. 10 (2024).
<https://doi.org/10.1007/s10462-024-10902-3>
- Boustani, Nada Mallah. "Artificial Intelligence Impact on Banks Clients and Employees in an Asian Developing Country." *Journal of Asia Business Studies* 16, no. 2 (2021): 267–278.
<https://doi.org/10.1108/jabs-09-2020-0376>
- Brodny, Jarosław, and Magdalena Tutak. "Analyzing the Level of Digitalization among the Enterprises of the European Union Member States and Their Impact on Economic Growth." *Journal of Open Innovation: Technology, Market, and Complexity* 8, no. 2 (2022): 70.
<https://doi.org/10.3390/joitmc8020070>
- Brodny, Jarosław, and Magdalena Tutak. "Assessing the Level of Digitalization and Robotization in the Enterprises of the European Union Member States." *PLOS ONE* 16, no. 7 (2021): e0254993.
<https://doi.org/10.1371/journal.pone.0254993>
- Brodny, Jarosław, and Magdalena Tutak. "Assessing the Level of Innovativeness and Digitalization of Enterprises in the European Union States." *Journal of Open Innovation: Technology, Market, and Complexity* 10, no. 1 (2024): 100210.
<https://doi.org/10.1016/j.joitmc.2024.100210>
- Burucu, Rukiye, and Hilal Türkben Polat. "Nursing Students' Readiness and Anxiety Regarding Medical Artificial Intelligence: A Mixed-methods Study." *Nurse Education in Practice* 88 (2025): 104538.
<https://doi.org/10.1016/j.nepr.2025.104538>
- Calabrese, Armando, Roberta Costa, Luigi Tiburzi, and Alexander Brem. "Merging Two Revolutions: A Human-artificial Intelligence Method to Study How Sustainability and Industry 4.0 are Intertwined." *Technological Forecasting and Social Change* 188 (2023):

122265.
<https://doi.org/10.1016/j.techfore.2022.122265>
- Capatina, Alexandru, Maher Kachour, Jessica Lichy, Adrian Micu, Angela-Eliza Micu, and Federica Codignola. "Matching the Future Capabilities of an Artificial Intelligence-based Software for Social Media Marketing with Potential Users' Expectations." *Technological Forecasting and Social Change* 151 (2020): 119794.
<https://doi.org/10.1016/j.techfore.2019.119794>
- Černevičienė, Jurgita, and Audrius Kabašinskas. "Explainable Artificial Intelligence (XAI) in Finance: A Systematic Literature Review." *Artificial Intelligence Review* 57, no. 8 (2024).
<https://doi.org/10.1007/s10462-024-10854-8>
- Çetin, Ensar. "Public Perceptions and Acceptance of Artificial Intelligence Humanoid Bots/ Robots: Evidence from Turkey." *Technological Forecasting and Social Change* 208 (2024): 123678.
<https://doi.org/10.1016/j.techfore.2024.123678>
- Chai, Tianlong, and Chengcheng Sha. "Prevention of AI Learning Anxiety: The Chain Mediation Role of Technophilia and AI Self-efficacy under the Context of Teacher Support." *Acta Psychologica* 261 (2025): 105979.
<https://doi.org/10.1016/j.actpsy.2025.105979>
- Chatterjee, Sheshadri, Ranjan Chaudhuri, Demetris Vrontis, Alkis Thrassou, and Soumya Kanti Ghosh. "Adoption of Artificial Intelligence-integrated CRM Systems in Agile Organizations in India." *Technological Forecasting and Social Change* 168 (2021): 120783.
<https://doi.org/10.1016/j.techfore.2021.120783>
- Chaturvedi, Rijul, Sanjeev Verma, Ronnie Das, and Yogesh K. Dwivedi. "Social Companionship with Artificial Intelligence: Recent Trends and Future Avenues." *Technological Forecasting and Social Change* 193 (2023): 122634.
<https://doi.org/10.1016/j.techfore.2023.122634>
- Chen, Xieling, Ziqing Liu, Li Wei, Jun Yan, Tianyong Hao, and Ruoyao Ding. "A Comparative Quantitative Study of Utilizing Artificial Intelligence on Electronic Health Records in the USA and China during 2008–2017." *BMC Medical Informatics and Decision Making* 18, no. S5 (2018).
<https://doi.org/10.1186/s12911-018-0692-9>
- Chen, Xinqiang, Jingying Ke, Xiu-e Zhang, and Jiangjie Chen. "The Impact of AI Anxiety on Employees' Work Passion: A Moderated Mediated Effect Model." *Acta Psychologica* 260 (2025): 105487.
<https://doi.org/10.1016/j.actpsy.2025.105487>
- Cheng, Can, Jiayu Luo, Chun Zhu, and Shangfeng Zhang. "Artificial Intelligence and the Skill Premium: A Numerical Analysis of Theoretical Models." *Technological Forecasting and Social Change* 200 (2024): 123140.
<https://doi.org/10.1016/j.techfore.2023.123140>
- Cheng, Eric Chi Keung, and Tianchong Wang. "Leading Digital Transformation and Eliminating Barriers for Teachers to Incorporate Artificial Intelligence in Basic Education in Hong Kong." *Computers and Education: Artificial Intelligence* 5 (2023): 100171.
<https://doi.org/10.1016/j.caeai.2023.100171>
- Cheng, Jiyang, Sunil Tiwari, Djebbouri Khaled, Mandeep Mahendru, and Umer Shahzad. "Forecasting Bitcoin Prices Using Artificial Intelligence: Combination of ML, SARIMA, and

-
- Facebook Prophet Models.” *Technological Forecasting and Social Change* 198 (2024): 122938.
<https://doi.org/10.1016/j.techfore.2023.122938>
- Cheung, L. H., P. Kodjamanova, A. Evans, and L. Mesana. “PNS77 Artificial Intelligence in Medical Devices: Regulatory and Reimbursement Landscape in the United States.” *Value in Health* 22 (2019): S299.
<https://doi.org/10.1016/j.jval.2019.04.1436>
- Chiu, Thomas Kin Fung, Qi Xia, Xinyan Zhou, Ching Sing Chai, and Miaoting Cheng. “Systematic Literature Review on Opportunities, Challenges, and Future Research Recommendations of Artificial Intelligence in Education.” *Computers and Education: Artificial Intelligence* 4 (2023): 100118.
<https://doi.org/10.1016/j.caeai.2022.100118>
- Cil, Abdullah Emir, and Kazim Yildiz. “A Systematic Literature Review on Applications of Explainable Artificial Intelligence in the Financial Sector.” *Internet of Things* 33 (2025): 101696.
<https://doi.org/10.1016/j.iot.2025.101696>
- Cintamür, İsmail Gökhan. “Acceptance of Artificial Intelligence Devices in Banking Services: Moderation Role of Technology Anxiety and Risk Aversion.” *International Journal of Bank Marketing* 42, no. 7 (2024): 2143–76.
<https://doi.org/10.1108/ijbm-10-2023-0563>
- Currie, Geoffrey, Johnathan Hewis, and Phillip Ebbs. “Gender Bias in Text-to-image Generative Artificial Intelligence Depiction of Australian Paramedics and First Responders.” *Australasian Emergency Care* 28, no. 2 (2025): 103–109.
<https://doi.org/10.1016/j.auec.2024.11.003>
- Dai, Yeming, Xinyu Yang, and Mingming Leng. “Forecasting Power Load: A Hybrid Forecasting Method with Intelligent Data Processing and Optimized Artificial Intelligence.” *Technological Forecasting and Social Change* 182 (2022): 121858.
<https://doi.org/10.1016/j.techfore.2022.121858>
- Dako, Farouk, Fabio Y. Moraes, Florence Doo, Mohammed Abed, Udunna Anazodo, Labenyimoh Patrick, Fernanda M. Favorito, Emadelden Fouad, Adewumi Alabi, Stephen Avery, Andrew Scott, Wilfred Ngwa, and Hedvig Hricak. “Digital Health and Artificial Intelligence Innovations for Oncology in Sub-Saharan Africa.” *The Lancet Oncology* 26, no. 10 (2025): e547–e557.
[https://doi.org/10.1016/S1470-2045\(25\)00360-2](https://doi.org/10.1016/S1470-2045(25)00360-2)
- Das, Sougato, and Biplab Datta. “Application of UTAUT2 on Adopting Artificial Intelligence Powered Lead Management System (AI-LMS) in Passenger Car Sales.” *Technological Forecasting and Social Change* 201 (2024): 123241.
<https://doi.org/10.1016/j.techfore.2024.123241>
- de la Paz-Marín, Mónica, Pilar Campoy-Muñoz, and César Hervás-Martínez. “Non-linear Multiclassifier Model Based on Artificial Intelligence to Predict Research and Development Performance in European Countries.” *Technological Forecasting and Social Change* 79, no. 9 (2012): 1731–1745.
<https://doi.org/10.1016/j.techfore.2012.06.001>
- Demir-Kaymak, Zeliha, Zekiye Turan, Nazli Unlu-Bidik, and Semiha Unkazan. “Effects of Midwifery and Nursing Students’ Readiness about Medical Artificial Intelligence on Artificial Intelligence Anxiety.” *Nurse Education in Practice* 78 (2024): 103994.
<https://doi.org/10.1016/j.nepr.2024.103994>

- Deng, Fengyi, Zichen Zhang, Yajun Lin, Xiaoping Mo, and Yue Zhang. "When Everyone Is Talking about AI: The Development of Fear of Missing out on AI Scale." *Telematics and Informatics* 100 (2025): 102283.
<https://doi.org/10.1016/j.tele.2025.102283>
- Di Vaio, Assunta, Rohail Hassan, and Claude Alavoine. "Data Intelligence and Analytics: A Bibliometric Analysis of Human–Artificial Intelligence in Public Sector Decision-Making Effectiveness." *Technological Forecasting and Social Change* 174 (2022): 121201.
<https://doi.org/10.1016/j.techfore.2021.121201>
- Dini, Niloofar, Saeed Yaghoubi, and Hamideh Bahrami. "Logistics Performance Index-Driven in Operational Planning for Logistics Companies: A Smart Transportation Approach." *Transport Policy* 160 (2025): 42–62.
<https://doi.org/10.1016/j.tranpol.2024.10.034>
- Durongkadej, Isarin, Wenyao Hu, and Heng Emily Wang. "How Artificial Intelligence Incidents Affect Banks and Financial Services Firms? A Study of Five Firms." *Finance Research Letters* 70 (2024): 106279.
<https://doi.org/10.1016/j.frl.2024.106279>
- Esteva, Andre, Alexandre Robicquet, Bharath Ramsundar, Volodymyr Kuleshov, Mark DePristo, Katherine Chou, Claire Cui, Greg Corrado, Sebastian Thrun, and Jeff Dean. "A Guide to Deep Learning in Healthcare." *Nature Medicine* 25, no. 1 (2019): 24–29.
<https://doi.org/10.1038/s41591-018-0316-z>
- Frank, Björn. "Consumer Preferences for Artificial Intelligence-enhanced Products: Differences Across Consumer Segments, Product Types, and Countries." *Technological Forecasting and Social Change* 209 (2024): 123774.
<https://doi.org/10.1016/j.techfore.2024.123774>
- Fu, Tong, Zhaoxuan Qiu, Xiangyang Yang, and Zijun Li. "The Impact of Artificial Intelligence on Green Technology Cycles in China." *Technological Forecasting and Social Change* 209 (2024): 123821.
<https://doi.org/10.1016/j.techfore.2024.123821>
- Garcés-Marín, Robinson, José Arias-Pérez, and Camilo Restrepo-Estrada. "The Interplay of Data-Driven Insights and AI Anxiety in Shaping the Impact of AI Capabilities on Circular Economy Capability." *Journal of Industrial Information Integration* 49 (2026): 101019.
<https://doi.org/10.1016/j.jii.2025.101019>
- Garcia, Pedro Pinheiro, Lucio Flavio Vismari, Joao Batista Camargo, Jorge Rady de Almeida, and Paulo Sergio Cugnasca. "Systematic Literature Review of Artificial Intelligence Techniques on Condition Based Maintenance Models for Transport Applications." *Engineering Applications of Artificial Intelligence* 164 (2026): 113230.
<https://doi.org/10.1016/j.engappai.2025.113230>
- Ghazwani, Salman, Patrick van Esch, Yuanyuan (Gina) Cui, and Prachi Gala. "Artificial Intelligence, Financial Anxiety and Cashier-Less Checkouts: A Saudi Arabian Perspective." *International Journal of Bank Marketing* 40, no. 6 (2022): 1200–1216.
<https://doi.org/10.1108/ijbm-09-2021-0444>
- Giwa, Fiyinfoluwa, and Sin-Yu Ho. "Artificial Intelligence Shock in Manufacturing: A Threat or an Opportunity for South Africa's Labour Market?" *Journal of Open Innovation: Technology, Market, and Complexity* 12, no. 1 (2026): 100696.
<https://doi.org/10.1016/j.joitmc.2025.100696>

-
- Gondwe, Gregory. "Artificial Intelligence, Journalism, and the Ubuntu Robot in Sub-Saharan Africa: Towards a Normative Framework." *Digital Journalism* (2024): 1–19.
<https://doi.org/10.1080/21670811.2024.2311258>
- Guggemos, Josef. "To Fear or Not to Fear – Human Resource Development Professionals' Positioning towards Artificial Intelligence with a Focus on Augmentation." *Computers and Education: Artificial Intelligence* 7 (2024): 100289.
<https://doi.org/10.1016/j.caeai.2024.100289>
- Gültekin, Vedat, and Ali Kavak. "An Assessment of Artificial Intelligence Anxieties of Academic Librarians in Türkiye." *The Journal of Academic Librarianship* 51, no. 3 (2025): 103058.
<https://doi.org/10.1016/j.acalib.2025.103058>
- Guo, Ge, and Yunpeng Wang. "An Integrated MPC and Deep Reinforcement Learning Approach to Trams-Priority Active Signal Control." *Control Engineering Practice* 110 (2021): 104758.
<https://doi.org/10.1016/j.conengprac.2021.104758>
- Gupta, Brij B., Akshat Gaurav, Prabin Kumar Panigrahi, and Varsha Arya. "Analysis of Artificial Intelligence-Based Technologies and Approaches on Sustainable Entrepreneurship." *Technological Forecasting and Social Change* 186 (2023): 122152.
<https://doi.org/10.1016/j.techfore.2022.122152>
- Ha, Taehyun. "An Explainable Artificial-intelligence-based Approach to Investigating Factors that Influence the Citation of Papers." *Technological Forecasting and Social Change* 184 (2022): 121974.
<https://doi.org/10.1016/j.techfore.2022.121974>
- Haque, AKM Bahalul, A.K.M. Najmul Islam, and Patrick Mikalef. "Explainable Artificial Intelligence (XAI) from a User Perspective: A Synthesis of Prior Literature and Problematising Avenues for Future Research." *Technological Forecasting and Social Change* 186 (2023): 122120.
<https://doi.org/10.1016/j.techfore.2022.122120>
- Hitsuwari, Jimpei, and Ryota Takano. "Associating Attitudes towards AI and Ambiguity: The Distinction of Acceptance and Fear of AI." *Acta Psychologica* 260 (2025): 105581.
<https://doi.org/10.1016/j.actpsy.2025.105581>
- Hoyos, William, Jose Aguilar, and Mauricio Toro. "Dengue Models Based on Machine Learning Techniques: A Systematic Literature Review." *Artificial Intelligence in Medicine* 119 (2021): 102157.
<https://doi.org/10.1016/j.artmed.2021.102157>
- Hu, Wenyi, Sanil Joseph, Rui Li, Ekaterina Woods, Jason Sun, Mingwang Shen, Catherine Lingxue Jan, Zhuoting Zhu, Mingguang He, and Lei Zhang. "Population Impact and Cost-Effectiveness of Artificial Intelligence-Based Diabetic Retinopathy Screening in People Living with Diabetes in Australia: A Cost Effectiveness Analysis." *EClinicalMedicine* 67 (2024): 102387.
<https://doi.org/10.1016/j.eclinm.2023.102387>
- Huang, GuoQiong Ivanka, IpKin Anthony Wong, Chen Jason Zhang, and Qilin Liang. "Generative AI Inspiration and Hotel Recommendation Acceptance: Does Anxiety over Lack of Transparency Matter?" *International Journal of Hospitality Management* 126 (2025): 104112.
<https://doi.org/10.1016/j.ijhm.2025.104112>

- Hughes, Jasmine G., Ketan Tamirisa, Ammar Dharani, Faraan Rahim, Erick Mahatara, Ntuli A. Kapologwe, and Almotasm Ben Hmeda. "Leveraging Artificial Intelligence to Enhance Oncology Care Delivery in Sub-Saharan Africa." *Mayo Clinic Proceedings: Digital Health* 4, no. 1 (2026): 100337.
<https://doi.org/10.1016/j.mcpdig.2026.100337>
- Ishii, Euma, Daniel K. Ebner, Satoshi Kimura, Louis Agha-Mir-Salim, Ryo Uchimido, and Leo A. Celi. "The Advent of Medical Artificial Intelligence: Lessons from the Japanese Approach." *Journal of Intensive Care* 8, no. 1 (2020).
<https://doi.org/10.1186/s40560-020-00452-5>
- Jabeur, Sami Ben, Cheima Gharib, Salma Mefteh-Wali, and Wissal Ben Arfi. "CatBoost Model and Artificial Intelligence Techniques for Corporate Failure Prediction." *Technological Forecasting and Social Change* 166 (2021): 120658.
<https://doi.org/10.1016/j.techfore.2021.120658>
- Jabeur, Sami Ben, Houssein Ballouk, Salma Mefteh-Wali, and Anis Omri. "Forecasting the Macrolevel Determinants of Entrepreneurial Opportunities Using Artificial Intelligence Models." *Technological Forecasting and Social Change* 175 (2022): 121353.
<https://doi.org/10.1016/j.techfore.2021.121353>
- Jacobs, Julian. "The Artificial Intelligence Shock and Socio-political Polarization." *Technological Forecasting and Social Change* 199 (2024): 123006.
<https://doi.org/10.1016/j.techfore.2023.123006>
- Jiang, Pengtao, Wanshu Niu, Qiaoli Wang, Ruizhi Yuan, and Keyu Chen. "Understanding Users' Acceptance of Artificial Intelligence Applications: A Literature Review." *Behavioral Sciences* 14, no. 8 (2024): 671.
<https://doi.org/10.3390/bs14080671>
- Jo, Hyeon, and Do-Hyung Park. "The Fear of Being Replaced by Generative AI: An Examination of Influential Factors among Office Workers." *Technological Forecasting and Social Change* 220 (2025): 124326.
<https://doi.org/10.1016/j.techfore.2025.124326>
- Johnson, Prince Chacko, Christofer Laurell, Mart Ots, and Christian Sandström. "Digital Innovation and the Effects of Artificial Intelligence on Firms' Research and Development – Automation or Augmentation, Exploration or Exploitation?" *Technological Forecasting and Social Change* 179 (2022): 121636.
<https://doi.org/10.1016/j.techfore.2022.121636>
- Kabudi, Tumaini, Ilias Pappas, and Dag Håkon Olsen. "AI-Enabled Adaptive Learning Systems: A Systematic Mapping of the Literature." *Computers and Education: Artificial Intelligence* 2 (2021): 100017.
<https://doi.org/10.1016/j.caeai.2021.100017>
- Kanzola, Anna-Maria, Konstantina Papaioannou, and Panagiotis Petrakis. "Unlocking Society's Standings in Artificial Intelligence." *Technological Forecasting and Social Change* 200 (2024): 123106.
<https://doi.org/10.1016/j.techfore.2023.123106>
- Kareem, Kola Yusuff, Yeonjeong Seong, Kyungtak Kim, and Younghun Jung. "A Case Study of Tidal Analysis Using Theory-Based Artificial Intelligence Techniques for Disaster Management in Taehwa River, South Korea." *Water* 14, no. 14 (2022): 2172.
<https://doi.org/10.3390/w14142172>

-
- Kommusaar, Janne, Silja Elunurm, Taridzo Chomutare, Mari Kangasniemi, Sanna Salanterä, and Laura-Maria Peltonen. "A Roadmap for Federated Learning Projects Using Health Data to Guide Sustainable Artificial Intelligence Development in the European Union." *International Journal of Medical Informatics* 208 (2026): 106242.
<https://doi.org/10.1016/j.ijmedinf.2025.106242>
- Koo, Bonhak, Catherine Curtis, and Bill Ryan. "Examining the Impact of Artificial Intelligence on Hotel Employees through Job Insecurity Perspectives." *International Journal of Hospitality Management* 95 (2021): 102763.
<https://doi.org/10.1016/j.ijhm.2020.102763>
- Labib, Ebtisam. "Artificial Intelligence in Marketing: Exploring Current and Future Trends." *Cogent Business and Management* 11, no. 1 (2024).
<https://doi.org/10.1080/23311975.2024.2348728>
- Labib, Ebtisam. "Artificial Intelligence in Marketing: Exploring Current and Future Trends." *Cogent Business & Management* 11, no. 1 (May 2, 2024).
<https://doi.org/10.1080/23311975.2024.2348728>
- Laupichler, Matthias Carl, Alexandra Aster, Jana Schirch, and Tobias Raupach. "Artificial Intelligence Literacy in Higher and Adult Education: A Scoping Literature Review." *Computers and Education: Artificial Intelligence* 3 (2022): 100101.
<https://doi.org/10.1016/j.caeai.2022.100101>
- Li, Chunxiao, Jiaqi Su, and Yufan Yang. "Understanding AI Anxiety Based on Terror Management Theory: A Meta Analytical Construction." *International Journal of Information Management* 88 (2026): 103043.
<https://doi.org/10.1016/j.ijinfomgt.2026.103043>
- Li, Jian, and Jin-Song Huang. "Dimensions of Artificial Intelligence Anxiety Based on the Integrated Fear Acquisition Theory." *Technology in Society* 63 (2020): 101410.
<https://doi.org/10.1016/j.techsoc.2020.101410>
- Li, Wei, Yaping Chang, Yinghao He, and Hong Li. "Research on Fear of Artificial Intelligence among the Elderly: The Key of Education and the Prison of Female." *Journal of Retailing and Consumer Services* 87 (2025): 104404.
<https://doi.org/10.1016/j.jretconser.2025.104404>
- Li, Wenyu, and Xueen Liu. "Anxiety about Artificial Intelligence from Patient and Doctor-physician." *Patient Education and Counseling* 133 (2025): 108619.
<https://doi.org/10.1016/j.pec.2024.108619>
- Liang, Yanlong, and Yun Zhai. "The Impact of Artificial Intelligence Impact Awareness on College Students' Employment Risk Perception: A Moderated Mediation Model." *Acta Psychologica* 261 (2025): 105808.
<https://doi.org/10.1016/j.actpsy.2025.105808>
- Liao, Chien-Hsiang, Yu-Hui Fang, and Chia-Ying Li. "Out of Gratification or Fears? A Dual Model to Explore Drivers of Generative Artificial Intelligence Adoption." *Technology in Society* 85 (2026): 103177.
<https://doi.org/10.1016/j.techsoc.2025.103177>
- Liu, Jun, Huihong Chang, Jeffrey Yi-Lin Forrest, and Baohua Yang. "Influence of Artificial Intelligence on Technological Innovation: Evidence from the Panel Data of China's Manufacturing Sectors." *Technological Forecasting and Social Change* 158 (2020): 120142.
<https://doi.org/10.1016/j.techfore.2020.120142>

- Liu, Peide, Hasan Dinçer, and Serhat Yüksel. "Multidimensional Assessment of the European Energy Union: Integrating Artificial Intelligence and Quantum Fuzzy Ranking Approaches." *Applied Soft Computing* 171 (2025): 112735.
<https://doi.org/10.1016/j.asoc.2025.112735>
- Liu, Yang, Zhenzhou Ying, Ying Ying, Ding Wang, and Jin Chen. "Artificial Intelligence Orientation and Internationalization Speed: A Knowledge Management Perspective." *Technological Forecasting and Social Change* 205 (2024): 123517.
<https://doi.org/10.1016/j.techfore.2024.123517>
- Lyu, Hui, Pengfei Shan, Chunli Hou, and Sufen Duan. "Artificial Intelligence for Student Performance Prediction in Blended Learning: A Systematic Literature Review." *Neurocomputing* 658 (2025): 131659.
<https://doi.org/10.1016/j.neucom.2025.131659>
- Magdin, Martin. "Literature Review: Current Trends and Advances in the Use of Artificial Intelligence for Ensuring the Safety and Efficiency of Gas Pipeline Operations." *Results in Engineering* 28 (2025): 107309.
<https://doi.org/10.1016/j.rineng.2025.107309>
- Malik, Nishtha, Shalini Nath Tripathi, Arpan Kumar Kar, and Shivam Gupta. "Impact of Artificial Intelligence on Employees Working in Industry 4.0 Led Organizations." *International Journal of Manpower* 43, no. 2 (2021): 334–54.
<https://doi.org/10.1108/ijm-03-2021-0173>
- Marda, Vidushi. "Artificial Intelligence Policy in India: A Framework for Engaging the Limits of Data-driven Decision-making." *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 376, no. 2133 (2018): 20180087.
<https://doi.org/10.1098/rsta.2018.0087>
- Mathenge, Wanjiku Ciku. "Artificial Intelligence for Diabetic Retinopathy Screening in Africa." *The Lancet Digital Health* 1, no. 1 (2019): e6–e7.
[https://doi.org/10.1016/s2589-7500\(19\)30009-3](https://doi.org/10.1016/s2589-7500(19)30009-3)
- Mbewe, Alinafe, Hammond N. Sarkwah, Olugbenga Ijaodola, Jirair Ratevosian, Solomon Simba, Tyler R. Smith, and Wisdom K. Atiwoto. "The Role of Digital Health and Artificial Intelligence in Improving the Reach and Effectiveness of HIV Prevention in Africa." *The Lancet Global Health* 14, no. 1 (2026): e131–42.
[https://doi.org/10.1016/s2214-109x\(25\)00447-4](https://doi.org/10.1016/s2214-109x(25)00447-4)
- Méndez-Suárez, Mariano, Maja Cukusic, and Ivana Ninčević-Pašalić. "AI FoMO (Fear of Missing Out) in the Workplace." *Technology in Society* 84 (2026): 103052.
<https://doi.org/10.1016/j.techsoc.2025.103052>
- Misra, Santosh K., Sujeet Kumar Sharma, Sumeet Gupta, and Satyasiba Das. "A Framework to Overcome Challenges to the Adoption of Artificial Intelligence in Indian Government Organizations." *Technological Forecasting and Social Change* 194 (2023): 122721.
<https://doi.org/10.1016/j.techfore.2023.122721>
- Montag, Christian, Johannes Kraus, Martin Baumann, and Dmitri Rozgonjuk. "The Propensity to Trust in (Automated) Technology Mediates the Links between Technology Self-Efficacy and Fear and Acceptance of Artificial Intelligence." *Computers in Human Behavior Reports* 11 (2023): 100315.
<https://doi.org/10.1016/j.chbr.2023.100315>

-
- Moravec, Vaclav, Nik Hynek, Marinko Skare, Beata Gavurova, and Matus Kubak. "Human or Machine? The Perception of Artificial Intelligence in Journalism, Its Socio-Economic Conditions, and Technological Developments toward the Digital Future." *Technological Forecasting and Social Change* 200 (2024): 123162.
<https://doi.org/10.1016/j.techfore.2023.123162>
- Muehlematter, Urs J., Paola Daniore, and Kerstin N. Vokinger. "Approval of Artificial Intelligence and Machine Learning-Based Medical Devices in the USA and Europe (2015–20): A Comparative Analysis." *The Lancet Digital Health* 3, no. 3 (2021): e195–203.
[https://doi.org/10.1016/s2589-7500\(20\)30292-2](https://doi.org/10.1016/s2589-7500(20)30292-2)
- Nassif, Ali Bou, Manar Abu Talib, Qassim Nasir, Yaman Afadar, and Omar Elgendy. "Breast Cancer Detection Using Artificial Intelligence Techniques: A Systematic Literature Review." *Artificial Intelligence in Medicine* 127 (2022): 102276.
<https://doi.org/10.1016/j.artmed.2022.102276>
- Nguyen, Tan Minh, Vito Carlo Alberto Caponio, Ravikumar Rajappa, Daniel Lawrence, Youstina Tawadros, Bree Jones, Keegan Crow, and Nicola Cirillo. "Face Validation of an Artificial Intelligence Driven Tool for Clinical Triage in Australian Public Oral Healthcare: A Pilot Study." *Value in Health*, (2025).
<https://doi.org/10.1016/j.jval.2025.12.001>
- Odugbesan, Jamiu Adetola, Sahar Aghazadeh, Rawan Enad Al Qaralleh, and Olukunle Samuel Sogeke. "Green Talent Management and Employees' Innovative Work Behavior: The Roles of Artificial Intelligence and Transformational Leadership." *Journal of Knowledge Management* 27, no. 3 (2022): 696–716.
<https://doi.org/10.1108/jkm-08-2021-0601>
- Özbek Güven, Gamze, Şerife Yılmaz, and Feyza Inceoğlu. "Determining Medical Students' Anxiety and Readiness Levels about Artificial Intelligence." *Heliyon* 10, no. 4 (2024): e25894.
<https://doi.org/10.1016/j.heliyon.2024.e25894>
- Park, Joonhyeong. "A Systematic Literature Review of Generative Artificial Intelligence (GenAI) Literacy in Schools." *Computers and Education: Artificial Intelligence* 9 (2025): 100487.
<https://doi.org/10.1016/j.caeai.2025.100487>
- Pedersen, Mangor, Heath R. Pardoe, Anton de Weger, Donna Hutchison, David F. Abbott, Karin Verspoor, and Graeme D. Jackson. "Toward Responsible Artificial Intelligence in Medicine: Reflections from the Australian Epilepsy Project." *Artificial Intelligence in Medicine* 167 (2025): 103192.
<https://doi.org/10.1016/j.artmed.2025.103192>
- Pineda, Vanessa García, Alejandro Valencia-Arias, Francisco Eugenio López Giraldo, and Edison Andrés Zapata-Ochoa. "Integrating Artificial Intelligence and Quantum Computing: A Systematic Literature Review of Features and Applications." *International Journal of Cognitive Computing in Engineering* 7 (2026): 26–39.
<https://doi.org/10.1016/j.ijcce.2025.08.002>
- Qian, Cheng, Chun Zhu, Duen-Huang Huang, and Shangfeng Zhang. "Examining the Influence Mechanism of Artificial Intelligence Development on Labor Income Share through Numerical Simulations." *Technological Forecasting and Social Change* 188 (2023): 122315.
<https://doi.org/10.1016/j.techfore.2022.122315>
- Qin, Weiwei. "How to Unleash Frugal Innovation through Internet of Things and Artificial Intelligence: Moderating Role of Entrepreneurial Knowledge and Future Challenges."

- Technological Forecasting and Social Change* 202 (2024): 123286.
<https://doi.org/10.1016/j.techfore.2024.123286>
- Qiu, Sha Sarah, Luxi Zhang, Fei You, and Xinshu Zhao. “Unpacking Media Channel Effects on AI Perception: A Network Analysis of AI Information Exposure across Channels, Overload, Literacy, and Anxiety among Chinese Users.” *Computers in Human Behavior* 173 (2025): 108790.
<https://doi.org/10.1016/j.chb.2025.108790>
- Rashidieranjbar, Fatemeh, Amirfarhad Farhadi, Azadeh Zamanifar, and Majid Sorouri. “A Systematic Literature Review: Generative Artificial Intelligence Applications for Ground Mobile Robot Navigation.” *Computers and Electrical Engineering* 130 (2026): 110906.
<https://doi.org/10.1016/j.compeleceng.2025.110906>
- Rehman, Ubaid ur. “The Future Role of Artificial Intelligence in Energy Management Systems for Smart Cities: A Systematic Literature Review of Trends, Gaps, and Future Direction.” *Sustainable Computing: Informatics and Systems* 49 (2026): 101249.
<https://doi.org/10.1016/j.suscom.2025.101249>
- Sapru, Aakash. “Psychological Resistance to AI: How Regulatory Focus Fuels AI Anxiety and Negative Attitudes toward AI.” *Technology in Society* 86 (2026): 103251.
<https://doi.org/10.1016/j.techsoc.2026.103251>
- Sarkar, Arkapravo. “From Anxiety to Advantage: Guiding Employees to embrace AI at Work.” *Organizational Dynamics* 54, no. 4 (2025): 101170.
<https://doi.org/10.1016/j.orgdyn.2025.101170>
- Sarmas, Elissaios, Alexandre Lucas, Andrés Acosta, Ferdinanda Ponci, Pedro Rodriguez, and Vangelis Marinakis. “A Federated Artificial Intelligence Testing and Experimentation Facility for the European Energy Sector.” *Engineering Applications of Artificial Intelligence* 164 (2026): 113400.
<https://doi.org/10.1016/j.engappai.2025.113400>
- Schiavo, Gianluca, Stefano Businaro, and Massimo Zancanaro. “Comprehension, Apprehension, and Acceptance: Understanding the Influence of Literacy and Anxiety on Acceptance of Artificial Intelligence.” *Technology in Society* 77 (2024): 102537.
<https://doi.org/10.1016/j.techsoc.2024.102537>
- Sezgin, Merve Gozde, and Hicran Bektas. “Artificial Intelligence-assisted Nursing in Cancer Care: A Meta-analysis of Its Impact on Pain, Anxiety, and Quality of Life.” *Applied Nursing Research* 84 (2025): 151987.
<https://doi.org/10.1016/j.apnr.2025.151987>
- Shen, Jie, Fanghao Zheng, Fiseha Berhanu Tesema, Wu Deng, Anthony Graham Bellotti, and Jing Xie. “Deep Learning and Explainable Artificial Intelligence for Large-Scale Photovoltaic Suitability Analysis in Australian Cities.” *Renewable Energy* 259 (2026): 125025.
<https://doi.org/10.1016/j.renene.2025.125025>
- Sikhakhane, Mbali. “Proficiency of Records Practitioners Using Artificial Intelligence to Enhance Records Management Practices at the Gauteng Department of Education in South Africa.” *Collection and Curation* 44, no. 3 (2025): 82–89.
<https://doi.org/10.1108/cc-08-2024-0039>
- Sjödin, David, Vinit Parida, and Marko Kohtamäki. “Artificial Intelligence Enabling Circular Business Model Innovation in Digital Servitization: Conceptualizing Dynamic Capabilities, AI Capacities, Business Models and Effects.” *Technological Forecasting and Social Change*

197 (2023): 122903.

<https://doi.org/10.1016/j.techfore.2023.122903>

Sohrabpour, Vahid, Pejvak Oghazi, Reza Toorajipour, and Ali Nazarpour. "Export Sales Forecasting Using Artificial Intelligence." *Technological Forecasting and Social Change* 163 (2021): 120480.

<https://doi.org/10.1016/j.techfore.2020.120480>

Soysal, Ganime Esra, Mehmet Ali CALIŞKAN, and Aykut TURGUT. "The Relationship between Anxiety about Artificial Intelligence and Nurses' Perceptions of Job Security: The Impact of Technological Transformation." *Applied Nursing Research* 86 (2025): 152023.

<https://doi.org/10.1016/j.apnr.2025.152023>

Standley, Claire J., J. Gabrielle Breugelmanns, Amol Chaudhari, Neil Cherian, Sabrina Chwalek, Arminster Deol, and Janan Dietrich. "Artificial Intelligence for Health Security in Africa: Benefits, Risks and Opportunities." *Epidemics* 53 (2025): 100870.

<https://doi.org/10.1016/j.epidem.2025.100870>

Teferi, Mikiyas G., Biruk T. Mengistie, Helina K. Teklehaimanot, Chernet T. Mengistie, Fitsum A. Gemechu, Michael A. Negussie, Tilahun J. Jufara, and Getaw W. Hassen. "Artificial Intelligence in Clinical Toxicology in Africa: Emerging Applications and Barriers." *African Journal of Emergency Medicine* 15, no. 4 (2025): 100901.

<https://doi.org/10.1016/j.afjem.2025.100901>

The Lancet Digital Health. "Can Artificial Intelligence Help Create Racial Equality in the USA?" *The Lancet Digital Health* 3, no. 3 (2021): e135.

[https://doi.org/10.1016/s2589-7500\(21\)00023-6](https://doi.org/10.1016/s2589-7500(21)00023-6)

Trocin, Cristina, Ingrid Våge Hovland, Patrick Mikalef, and Christian Dremel. "How Artificial Intelligence Affords Digital Innovation: A Cross-Case Analysis of Scandinavian Companies." *Technological Forecasting and Social Change* 173 (2021): 121081.

<https://doi.org/10.1016/j.techfore.2021.121081>

Truong, Yann, and Savvas Papagiannidis. "Artificial Intelligence as an Enabler for Innovation: A Review and Future Research Agenda." *Technological Forecasting and Social Change* 183 (2022): 121852.

<https://doi.org/10.1016/j.techfore.2022.121852>

Tutar, Şerife, Hande Özgörü, and Züleyha Ögür. "Determining Pediatric Nurses' Anxiety Levels, Concerns, and Metaphor Perceptions towards Artificial Intelligence Technologies: A Mixed-Method Study." *Journal of Pediatric Nursing* 85 (2025): 441–50.

<https://doi.org/10.1016/j.pedn.2025.09.005>

Ünal, Ayşe Sevim, and Aydın Avcı. "Evaluation of Neonatal Nurses' Anxiety and Readiness Levels towards the Use of Artificial Intelligence." *Journal of Pediatric Nursing* 79 (2024): e16–23.

<https://doi.org/10.1016/j.pedn.2024.09.012>

Urom, Christian, Gideon Ndubuisi, Khaled Guesmi, and Ramzi Benkraien. "Quantile Co-Movement and Dependence between Energy-Focused Sectors and Artificial Intelligence." *Technological Forecasting and Social Change* 183 (2022): 121842.

<https://doi.org/10.1016/j.techfore.2022.121842>

Vasile, Mirela, Maria-Laura Tîrlă, Juha Lemmetyinen, Răzvan Popescu, Toma Hoinic, and Ionuţ Şandric. "Detection of Thermokarst Landforms in the European Arctic Region Using Earth Observation and Artificial Intelligence." *Remote Sensing Applications: Society and*

- Environment*, (2025): 101832.
<https://doi.org/10.1016/j.rsase.2025.101832>
- Vasilescu, Maria Denisa, Andreea Claudia Serban, Gina Cristina Dimian, Mirela Ionela Aceleanu, and Xose Picatoste. “Digital Divide, Skills and Perceptions on Digitalisation in the European Union—Towards a Smart Labour Market.” Edited by Javier Ordóñez-Monfort. *PLOS ONE* 15, no. 4 (2020): e0232032.
<https://doi.org/10.1371/journal.pone.0232032>
- Vásquez, Óscar, Francesco D. Sandulli, and Jorge Gallego. “Technological Agglomerations and the Emergence of Artificial Intelligence Start-up Ecosystems in Europe.” *Technology in Society* 84 (2026): 103092.
<https://doi.org/10.1016/j.techsoc.2025.103092>
- Verma, Surabhi, and Vibhav Singh. “Impact of Artificial Intelligence-Enabled Job Characteristics and Perceived Substitution Crisis on Innovative Work Behavior of Employees from High-Tech Firms.” *Computers in Human Behavior* 131 (2022): 107215.
<https://doi.org/10.1016/j.chb.2022.107215>
- Verma, Surabhi, and Vibhav Singh. “The Employees Intention to Work in Artificial Intelligence-Based Hybrid Environments.” *IEEE Transactions on Engineering Management* 71 (2024): 3266–3277.
<https://doi.org/10.1109/tem.2022.3193664>
- Vicente, María Rosalía, Carla Álvarez-Rodríguez, and Ana Suárez-Álvarez. “An Old Familiar Song? Assessing the Artificial Intelligence Divide among the Regions of the European Union and Its Connections with Digital Divides.” *Telecommunications Policy* 49, no. 8 (2025): 103030.
<https://doi.org/10.1016/j.telpol.2025.103030>
- Vo, Vinh, Maame E. Woode, Stacy M. Carter, Chris Degeling, and Gang Chen. “Unravelling Public Preferences for the Use of Artificial Intelligence Mobile Health Applications in Australia.” *Value in Health* 28, no. 11 (2025): 1696–1704.
<https://doi.org/10.1016/j.jval.2025.06.013>
- Volkmar, Gioia, Peter M. Fischer, and Sven Reinecke. “Artificial Intelligence and Machine Learning: Exploring Drivers, Barriers, and Future Developments in Marketing Management.” *Journal of Business Research* 149 (2022): 599–614.
<https://doi.org/10.1016/j.jbusres.2022.04.007>
- Wang, Canghong, Raima Nazar, Sajid Ali, and Muhammad Saeed Meo. “Eco-Friendly Algorithms: Artificial Intelligence and Green Finance in European Intelligent Nations.” *Technology in Society* 83 (2025): 102960.
<https://doi.org/10.1016/j.techsoc.2025.102960>
- Wang, Xinggang, and Yuhua Wen. “The Role of Artificial Intelligence Tools on Chinese EFL Learners’ Speaking Skills, Speaking Anxiety, and Motivation.” *Acta Psychologica* 261 (2025): 105957.
<https://doi.org/10.1016/j.actpsy.2025.105957>
- Wang, Yan, and Shuai Xu. “Relationship between Artificial Intelligence Tool Usage Experience and Academic Stress among College Students: Mediating Role of Loneliness and Moderating Role of Academic Self-efficacy.” *Acta Psychologica* 263 (2026): 106220.
<https://doi.org/10.1016/j.actpsy.2026.106220>

-
- Xu, Yingwei (Wayne), Gongmei (May) Zhou, Ruiying (Raine) Cai, and Dogan Gursoy. "When Disclosing the Artificial Intelligence (AI) Technology Integration into Service Delivery Backfires: Roles of Fear of AI, Identity Threat and Existential Threat." *International Journal of Hospitality Management* 122 (2024): 103829.
<https://doi.org/10.1016/j.ijhm.2024.103829>
- Yang, Jiaqi, Yvette Blount, and Alireza Amrollahi. "Artificial Intelligence Adoption in a Professional Service Industry: A Multiple Case Study." *Technological Forecasting and Social Change* 201 (2024): 123251.
<https://doi.org/10.1016/j.techfore.2024.123251>
- Yang, Ronghui. "Between Epistemic Empowerment and Moral Anxiety: Chinese Patients' Ambivalence toward AI-Assisted Diagnosis." *Social Science & Medicine* 391 (2026): 118929.
<https://doi.org/10.1016/j.socscimed.2026.118929>
- Young, Sean D, Jeffrey S Crowley, and Sten H Vermund. "Artificial Intelligence and Sexual Health in the USA." *The Lancet Digital Health* 3, no. 8 (2021): e467–68.
[https://doi.org/10.1016/s2589-7500\(21\)00117-5](https://doi.org/10.1016/s2589-7500(21)00117-5)
- Yu, Chi-Lin. "Fear of Missing Out on AI: A Psychological Cost of Technological Revolution." *Telematics and Informatics Reports* 20 (2025): 100277.
<https://doi.org/10.1016/j.teler.2025.100277>
- Zhang, Cong, Yiwen Meng, and Xinyu Ma. "Artificial Intelligence in EFL Speaking: Impact on Enjoyment, Anxiety, and Willingness to Communicate." *System* 121 (2024): 103259.
<https://doi.org/10.1016/j.system.2024.103259>
- Zhang, Haoyang, and Luokuan Zhang. "How Design Ability Anxiety Leads to Design Students' Dependence on Artificial Intelligence? Internship and Professional Identity, Feedback and Evaluation Mechanisms, and the Role of Innovative Mindset." *Acta Psychologica* 259 (2025): 105423.
<https://doi.org/10.1016/j.actpsy.2025.105423>
- Zhang, Shengtai, Pengli Guo, Yiwei Yuan, and Yajun Ji. "Anxiety or Engaged? Research on the Impact of Technostress on Employees' Innovative Behavior in the Era of Artificial Intelligence." *Acta Psychologica* 259 (2025): 105442.
<https://doi.org/10.1016/j.actpsy.2025.105442>
- Zhao, Hairong, Bocong Yuan, and Yang Song. "Employees' Perception of Generative Artificial Intelligence and the Dark Side of Work Outcomes." *Journal of Hospitality and Tourism Management* 61 (2024): 191–99.
<https://doi.org/10.1016/j.jhtm.2024.10.007>
- Zhao, Peiya, Yu Gao, and Xue Sun. "How Does Artificial Intelligence Affect Green Economic Growth?—Evidence from China." *Science of The Total Environment* 834 (2022): 155306.
<https://doi.org/10.1016/j.scitotenv.2022.155306>
- Zhu, Jun, Jingting Zhang, and Yiqing Feng. "Hard Budget Constraints and Artificial Intelligence Technology." *Technological Forecasting and Social Change* 183 (2022): 121889.
<https://doi.org/10.1016/j.techfore.2022.121889>