Finanziato dall'Unione europea NextGenerationEU

MUR Ministero dell'Università e della Ricerca

Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA

Università di Foggia

Corso di Dottorato in
Scienze Economiche
XXXVIII ciclo

Beyond the Algorithm: Exploring Risk, Trust, and Value in the Adoption of AI-Driven Clinical Decision Support Systems

Tutor
Prof. Claudio Nigro

CLAUDIO NIGRO
07.04.2026
11:29:06
GMT+02:00

Dottoranda
Simona Curiello

Co-tutor
Prof.ssa Enrica Iannuzzi

ENRICA
IANNUZZI
08.04.2026
09:33:23
GMT+02:00

Anno Accademico 2024/2025

> “As algorithms push humans out of the job market, wealth and power might become concentrated in the hands of the tiny elite that owns the all-powerful algorithms.”
>
> – Yuval Noah Harari, Homo Deus (2016, p. 382)

Index

List of Abbreviations

The following abbreviations are used throughout this thesis for clarity and consistency.

Acronym	Meaning
AI	Artificial Intelligence
AI-CDSS	Artificial Intelligence-based Clinical Decision Support System
AIA	Artificial Intelligence Act
BDA	Big Data Analytics
CAGR	Compound Annual Growth Rate
CDSS	Clinical Decision Support System
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CI	Confidence Interval
CIM	Corporate Innovation Matrix
CNIPA	China National Intellectual Property Administration
CR	Composite Reliability
CSET	Center for Security and Emerging Technology
DCE	Discrete Choice Experiment
DOSPERT	Domani-specific Risk Taking
DTM	Document-Term Matrix
DV	Dependent Variable
EC	European Commission
EDF	Effective Degrees of Freedom
EE	Effort Expectancy
EHDS	European Health Data Space
EI	European Identity
EIT	European Institute of Innovation & Technology
EMA	European Medicines Agency
ENISA	European Union Agency for Cybersecurity
EPO	European Patent Office
ERGO	Ethics and Research Governance Online
EU	European Union
EUT	Expected Utility Theory
FC	Facilitating Conditions

FDA	Food and Drug Administration
FEC	Faculty Ethics Committee
GAAIS	General Attitudes towards Artificial Intelligence Scale
GAM	Generalized Additive Model
GDP	Gross Domestic Product
GDPR	General Data Protection Regulation
HCI	Human-Computer Interaction
HIT	Health Information Technology
ICT	Information and Communication Technology
IP	Intellectual Property
IPC	International Patent Classification
JPO	Japan Patent Office
KIPO	Korean Intellectual Property Office
KM	Knowledge Management
LDA	Latent Dirichlet Allocation
MHRA	Medicines and Healthcare Products Regulatory Agency
MIT	Massachusetts Institute of Technology
ML	Machine Learning
MLR	Robust Maximum Likelihood
MM	Motivational Model
NHS	National Health Service
NICE	National Institute for Health and Care Excellence
NLP	Natural Language Processing
OECD	Organization for Economic Co-operation and Development
OLS	Ordinary Least Square
PCB	Perceived Communication Barriers
PE	Performance Expectancy
PI	Professional Indemnity
PLI	Perceived Liability Issues
PMT	Perceived Mistrust
PPA	Perceived Performance Anxiety
PPC	Perceived Privacy Concerns
PR	Perceived Risks

PRISMA	Preferred Reporting Items for Systematic Review and Meta-Analysis
PSB	Perceived Social Biases
PT	Prospect Theory
PUS	Perceived Unregulated Standards
RMSEA	Root Mean Square Error of Approximation
SD	Standard Deviation
SEM	Structural Equation Modelling
SI	Social Influence
SLR	Systematic Literature Review
TA	Training Adequacy
TAM	Technology Acceptance Model
TC	Times Cited
TLI	Tucker–Lewis Index
TOPSIS	Technique for Order Preference by Similarity to an Idel Solution
TRI	Technology Readiness Index
UK	United Kingdom
UNESCO	United Nations Educational, Scientific and Cultural Organization
UPC	U.S. Patent Classification
USA	United States of America
USPTO	United States Patent and Trademark Office
UTAUT	Unified Theory of Acceptance and Use of Technology
VIF	Variance Inflation Rate
WHO	World Health Organization
WIPO	World Intellectual Property Organization
WTA	Willingness to Accept
WTP	Willingness to Pay
XAI	Explainable AI

Abstract

The advent of Artificial Intelligence (AI) within the healthcare sector represents one of the most revolutionary developments of the past few decades. Despite its potential to enhance diagnostic accuracy, treatment plans, and organizational efficiencies, the adoption of AI-powered Clinical Decision Support Systems (AI-CDSSs) remains low. This doctoral research aims at exploring the behavioural, organizational, and institutional issues that influence adoption, trust, and purchasing of AI solutions in healthcare, combining perspectives lenses from behavioural economics, innovation, and governance research streams.

Theoretically grounded in Prospect Theory (Kahneman & Tversky, 1979), this work investigates how perceptions of risk, uncertainty, and trust influence decision-making processes among clinicians and institutions. To understand this complex phenomenon at various levels of analysis, a multi-method research design was implemented.

The initial study offers a systematic review of the literature with the aim of clarifying the impact of cognitive biases, risk attitudes, and trust perceptions on the implementation of AI in the clinical setting. The second piece of work looks at the European Commission's official communications to examine the main policy narratives – human oversight, risk classification, and accountability – that characterize the changing governance of "trustworthy AI" in the healthcare sector.

In the third study, over 8,000 AI-healthcare patents (2014-2024) were subjected to patentometric and network analysis to locate worldwide innovation trajectories and thus demonstrate the central role of diagnostics, predictive analytics, and medical imaging ecosystems.

The fourth piece of this work consists of a cross-sectional survey of healthcare professionals in Italy and the UK. This is based on the Unified Theory of Acceptance and Use of Technology (UTAUT-Venkatesh, 2003) to identify the psychological and contextual factors that drive AI adoption.

The final study (*ongoing*) employs a Discrete Choice Experiment (DCE) with NHS procurement specialists to understand institutional decision-making and trade-offs between technical performance, fairness, and regulatory compliance.

Overall, the findings highlight that AI adoption in healthcare is influenced by a multitude of factors, including technical, behavioural and managerial dimensions. Additionally, this research offers actionable insightful perspectives for managers and policymakers seeking to align innovation, accountability and ethical standards in the implementation of AI in healthcare.

Introduction

The term Artificial Intelligence (AI) was coined in 1956 by Marvin Minsky and John McCarthy, computer scientists at MIT in Boston. However, the origins of AI can be traced back to the work conducted during World War II by Alan Turing. Despite substantial investments in the 1960s and beyond, AI successes were limited, leading to what has been termed the "AI winter" (Haenlein & Kaplan, 2019). Over the past two decades, however, rapid advances in computational capacity and the simultaneous availability of an unprecedented volume of data have progressively made AI a field of interest across numerous disciplines, including healthcare (Haenlein and Kaplan, 2019; Locatelli *et al*., 2021). Across various sectors in business and society, a prominent and widespread rise in the autonomy of technologies and algorithms can now be observed (Demetis & Lee, 2018). This trend is especially noticeable in discussions concerning the applications of Generative AI, including ChatGPT, which have raised public and scholarly attention to the transformative power of machine learning (ML) systems (Freisinger & Schneider, 2024). In the health sector, this technological momentum has become a defining feature of the current fourth industrial revolution, broadly termed as *Health 4.0*, a paradigm shift in which clinical and operational procedures increasingly incorporate robotics, data analytics, and intelligent systems (Schwab, 2016; Thuemmler & Bai, 2017). Enhanced computational capabilities now allow AI-driven technologies to perform complex tasks that were traditionally handled by humans (Logg *et al*., 2019).

Within this context, AI in healthcare can be broadly defined as the ability of technological systems to mimic human intelligence, specifically the ability to accurately interpret external data, learn from them, and apply learned knowledge to carry out specific tasks or assist humans in decision-making (Haenlein & Kaplan, 2019; Topol, 2019). Applications of AI have the potential to "*significantly modify diagnostic and therapeutic pathways, the decision-making processes of physicians, and ultimately, even the doctor-patient relationship*" (Locatelli *et al*., 2021, p. 37), potentially limiting the discretion of healthcare professionals (Giest & Klievink, 2022).

At the same time, the global ageing population and the steady increase in life expectancy are exerting growing pressure on healthcare systems worldwide, thus creating a pressing need to improve the quality, accessibility, and efficiency of care (World Health Organization [WHO], 2021). According to the most recent United Nations estimates (United Nations, 2024), the prevalence of chronic and multimorbid conditions, including neurological, respiratory, and cardiovascular diseases, will rise sharply by 2030. Nearly one in six people worldwide will be 65 years of age or older, and this figure is predicted to rise to one in four in North America and Europe (United Nations, 2024). Significant human and financial resources are needed to manage these complex patients, as well as integrated data infrastructures that can facilitate home-based and long-term care (OECD, 2021).

Across Europe, health expenditure is about 10% of the gross domestic product (GDP), ranging from about 7-8% in some Eastern states to over 11% in major Western states (Eurostat, 2024). Forecasts suggest the *health-spending-to-GDP ratio* will keep on going up during the next ten years driven by ageing population, chronic diseases burdens, and growing investment in digital/AI-enabled health systems (OECD & European Commission, 2024).

Additionally, the shortage of healthcare professionals is still a very serious problem. According to the WHO (2020), the global shortage of healthcare workers – including physicians, nurses, and midwives – will exceed around 10-11 million by 2030. Such a mix of increasing demand and limited supply highlights the necessity of digital innovation and technology as a means of alleviating the pressure on health services (Senbekov *et al*., 2020).

One of the most promising technologies to assist in addressing these systemic issues is AI. The European Parliament (2020) defines AI as "*the ability of a computer program to perform tasks or processes that we usually associate with human intelligence*". This technology has the potential to improve clinical decision-making, optimize administrative operations, and increase the precision and effectiveness of diagnostic procedures in the healthcare industry (Topol, 2019; Jiang *et al*., 2017). When used properly, AI can assist personalized treatment pathways, reduce routine workload, and enable real-time monitoring, freeing up clinicians to concentrate more on patient care (Rajkomar *et al*., 2019; Yu *et al*., 2018).

The greatest potential of AI to change the world is essentially the capability to learn from large and complicated datasets. Present-day systems use machine learning (ML), deep learning (DL), and natural language processing (NLP) to identify patterns, predict results, and extract insights that would be quite unreachable even by the most advanced traditional statistical methods (Esteva *et al*., 2019; Beam & Kohane, 2018). AI-driven imaging technologies are a significant example of systems that can now perform visual recognition tasks beyond human ability, attaining an error rate of less than 3 percent in diagnostic classification (Liu *et al*., 2019). Similarly, language models that have been trained with clinical text are becoming more accurate in performing automated reporting, triage, and risk prediction (Johnson *et al*., 2022).

However, the implementation of AI in the medical field is still at a very low level despite such developments. According to a WHO survey of the European Region (50 Member States, 2024-2025), the use of AI in healthcare is still very limited (WHO, 2025). For instance, less than 10% of EU-funded AI-health-care projects had reached clinical practice, and only a small percentage of nations had established fully operational-process AI systems. The problems standing in the way of this phenomenon are scattered datasets, regulatory uncertainty, lack of interoperability, and poor digital literacy among clinical staff (EIT Health & McKinsey & Company, 2020). Solving these problems

would mean that we need to put our money into data infrastructures that can work together, algorithms that are transparent and thorough change-management strategies that engage clinicians not only in the design but also in the deployment phases (Amann *et al*., 2020; Kelly *et al*., 2019).

Moreover, the choice to use AI in health care practice successfully hinges largely on trust, which can be achieved through explainability, robust evaluation, and consistence with moral and legal frameworks (Floridi & Chiriatti, 2020): interpretable, evidence-based, and user-centred systems are more likely to support clinical acceptance and long-term sustainability (Shortliffe & Sepúlveda, 2018).

Additionally, accountability and liability issues make the use of AI technologies even more difficult to be accepted in the clinical environment. In case a patient gets harmed because of an erroneous algorithmic recommendation, establishing responsibility among clinicians, developers, and healthcare institutions both from a legal and ethical point of view can become very complicated (Price *et al*., 2019). Hence, these challenges highlight the importance of robust regulatory frameworks – such as those outlined in the context of the European Union's Artificial Intelligence Act (EU AI Act – 2024/1689). These frameworks define in detail the aspects of human control, risk classification, and conformity standards for AI applications in the healthcare sector (European Parliament & Council of the European Union, 2024).

The thesis aims to provide a comprehensive framework for understanding the socio-technical and behavioural factors that exert influence on clinicians and institutions with regard to the assessment, trust, and ultimately selection of AI-CDSSs.[1] The intricate and multidisciplinary nature of this doctoral study necessitated the integration of several complementary fields of study, including decision sciences, innovation management, behavioural economics, and digital health policy. This multifaceted approach allowed to capture the intrinsic complexity inherent in the study of AI development, adoption and procurement in healthcare, wherein organizational, human, and technological factors interact in real-time.

This study makes both a conceptual and methodological contribution. From a conceptual standpoint, it addresses a significant gap in literature by advancing knowledge of the socio-technical and behavioural factors that influence clinicians' trust and willingness to use and procure AI in healthcare. This is due to the fact that prior research has predominantly focused on technical performance as opposed to organizational and human adoption dynamics. From a methodological standpoint, the

[1] Clinical Decision Support Systems (CDSSs) are computer-based tools that integrate evidence-based recommendations, alerts, or diagnostic support into electronic health records to assist healthcare professionals in making well-informed clinical decisions (Sutton *et al*., 2020). When enhanced with AI, these systems can analyse large, complicated datasets and provide tailored, adaptive advice, helping doctors with diagnosis, treatment planning, and resource allocation (Shortliffe & Sepúlveda, 2018).

combination of different but complementary techniques – systematic literature mapping, content analysis, patentometric analysis, cross-sectional survey modeling, and discrete choice experimentation – provides a novel framework for identifying adoption determinants from various perspectives and levels of analysis.

Finally, the thesis adds to the continuing discussions about governance and policy related to the trustworthy implementation of AI in healthcare systems. By grounding empirical findings in the context of European digital health strategies and regulatory frameworks, this work aligns with current policy trends to prioritize human-centered, ethical, and explicable AI.

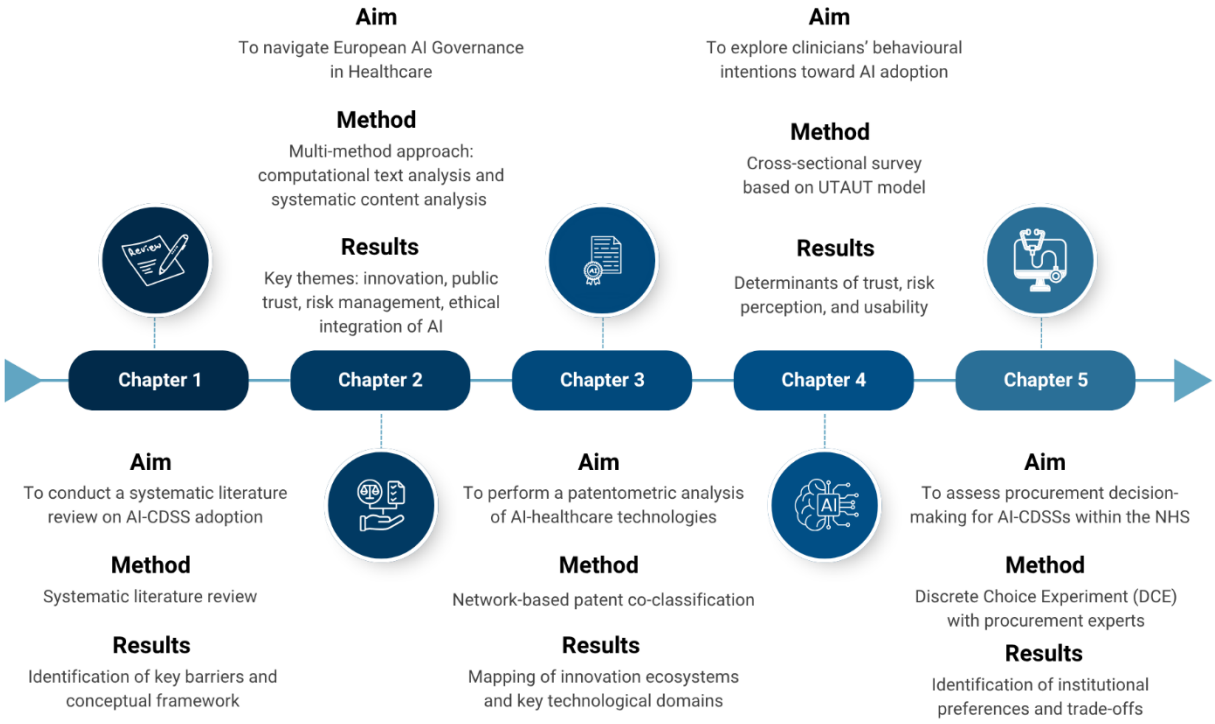
To address the aforementioned objectives, the present thesis consists of four chapters that together explore the phenomenon from *theoretical*, *technological*, *behavioural* and *institutional* perspectives:

- The first paper (*Chapter 1*) reports on a systematic review of the literature that seeks to clarify how medical practitioners perceive the concepts of risk, uncertainty, and trust when adopting AI-CDSSs. The work is based on Prospect Theory (Kahneman & Tversky, 1979) as a theoretical framework. Although AI's clinical benefits are consistently demonstrated, the results indicate that issues of trust, risk aversion, and cognitive biases are still major obstacles to implementation.
- The second piece of work (*Chapter 2*) undertakes a comprehensive examination of the European Commission's formal communications on AI in healthcare, combining both computational text analysis techniques and a systematic content analysis of 95 official documents published between 2018 and 2025. The findings indicate that the Commission's discourse places significant emphasis on themes such as innovation, public trust, risk management, and the ethical integration of AI.
- The third study (*Chapter 3*) carries out a worldwide patentometric analysis of over 8,000 AI-healthcare patents (2014-2024) to present a macro-level perspective of the technological innovations. The study makes use of network-based patent co-classification and centrality metrics to reveal that high-value patents are usually deeply rooted in highly interconnected technological ecosystems. These are dominated by the areas of medical imaging, diagnostics, and predictive analytics, thus pointing to the strategic innovation and knowledge diffusion spheres.
- The fourth work (*Chapter 4*) constitutes a cross-sectional survey of healthcare workers in Italy and the UK, who were invited to complete an extended Unified Theory of Acceptance and Use of Technology (UTAUT) model (Venkatesh *et al*., 2003). The analysis explores the potential influence of perceived risks, perceived benefits, and system usability on clinicians'

behavioural intentions to adopt AI-CDSSs. The research identifies psychological and contextual drivers for the adoption of such systems.

- The last research (*Chapter 5*) employs a Discrete Choice Experiment (DCE) with NHS procurement specialists to uncover the way in which institutional decision-makers weigh AI-CDSS key features such as workflow integration, algorithmic fairness, diagnostic accuracy, and regulatory robustness, in their purchasing decisions (Marshall *et al.*, 2010). The supplementary think-aloud interviews allow the participants to explain their thoughts and to identify the contextual factors behind the trade-offs they made.

Figure 1 – The Flow of Thesis Implementation



Source: author's own elaboration

Collectively, these five studies offer an integrated understanding of how clinicians and healthcare procurement actors assess and prioritize the value of AI-CDSS technologies. The findings inform the design of human-centred training strategies, evidence-based policy frameworks, and procurement approaches that align technological innovation with clinical safety, ethical standards, and operational trust within modern healthcare systems

References

Amann, J., Blasimme, A., Vayena, E., et al. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, 20, 310. https://doi.org/10.1186/s12911-020-01332-6

Beam, A. L., & Kohane, I. S. (2018). Big Data and Machine Learning in Health Care. *JAMA*, *319*(13), 1317–1318. https://doi.org/10.1001/jama.2017.18391

Demetis, D., & Lee, A. (2018). When Humans Using the IT Artifact Becomes IT Using the Human Artifact. *Journal of the Association for Information Systems*, *19*, 929–952. https://doi.org/10.17705/1jais.00514

EIT Health & McKinsey & Company. (2020). *Transforming healthcare with AI: The impact on the workforce and organisations*. EIT Health. https://eithealth.eu/wp-content/uploads/2020/03/EIT-Health-and-McKinsey_Transforming-healthcare-with-AI.pdf

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. *Nature medicine*, *25*(1), 24-29. https://doi.org/10.1038/s41591-018-0316-z

European Parliament & Council of the European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council on harmonised rules on artificial intelligence (Artificial Intelligence Act).* Official Journal of the European Union, L 1689. https://data.europa.eu/eli/reg/2024/1689/oj

Eurostat. (2024). *Healthcare expenditure statistics – overview*. European Commission. https://ec.europa.eu/eurostat/statistics-explained/index.php/Healthcare_expenditure_statistics_-_overview

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds and Machines*, *30*, 1–14. https://doi.org/10.1007/s11023-020-09548-1

Freisinger, E., & Schneider, S. (2024). Decoding decision delegation to artificial intelligence: A mixed-methods study on the preferences of decision-makers and decision-affected in surrogate decision contexts. *European Management Journal*. https://doi.org/10.1016/j.emj.2024.10.004

Giest, S., & Klievink, B. (2022). More than a digital system: How AI is changing the role of bureaucrats in different organizational contexts. *Public Management Review*, *26*, 1–20. https://doi.org/10.1080/14719037.2022.2095001

Haenlein, M., & Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. California Management Review, 61, 5-14. https://doi.org/10.1177/0008125619864925

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology, 2*(4), 230–243. https://doi.org/10.1136/svn-2017-000101

Johnson, M., Albizri, A., Harfouche, A., & Fosso-Wamba, S. (2022). Integrating human knowledge into artificial intelligence for complex and ill-structured problems: Informed artificial intelligence. *International Journal of Information Management*, *64*, 102479. https://doi.org/10.1016/j.ijinfomgt.2022.102479

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10.2307/1914185

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC medicine*, *17*(1), 195. https://doi.org/10.1186/s12916-019-1426-2

Liu, P., Lyu, M., King, I., & Xu, J. (2019). *SelFlow: Self-Supervised Learning of Optical Flow* (p. 4575). https://doi.org/10.1109/CVPR.2019.00470

Locatelli, R., Uselli, A., Pepe, G., & Salis, F. (2023). Artificial Intelligence: new data and new models in credit risk management. *RISK MANAGEMENT MAGAZINE*, *18*(3), 2-15. *https://dx.doi.org/10.47473/2020rmm0130*

Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behaviour and Human Decision Processes*, *151*, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005

OECD & European Commission. (2024). *Health at a Glance: Europe 2024 – State of Health in the EU Cycle*. OECD Publishing. https://doi.org/10.1787/b3704e14-en

Organisation for Economic Co-operation and Development. (2021). *Health at a Glance 2021: OECD Indicators* (OECD Publishing, Paris). https://doi.org/10.1787/ae3016b9-en

Price, W. N., Gerke, S. Cohen, G. (2019). Potential Liability for Physicians Using Artificial Intelligence. *AMA: Journal of the American Medical Association* 322, 18: 1765-1766. https://doi.org/10.1001/jama.2019.15064

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. *The New England journal of medicine*, 380(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259

Schwab, K. (2016). *The Fourth Industrial Revolution*. World Economic Forum.

Senbekov, M., Saliev, T., Bukeyeva, Z., Almabayeva, A., Zhanaliyeva, M., Aitenova, N., Toishibekov, Y., & Fakhradiyev, I. (2020). The Recent Progress and Applications of Digital Technologies in Healthcare: A Review. *International journal of telemedicine and applications*, *2020*, 8830200. https://doi.org/10.1155/2020/8830200

Shortliffe, E. H., & Sepúlveda, M. J. (2018). Clinical Decision Support in the Era of Artificial Intelligence. *JAMA*, *320*(21), 2199–2200. https://doi.org/10.1001/jama.2018.17163

Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. *NPJ Digital Medicine, 3*, 17. https://doi.org/10.1038/s41746-020-0221-y

Thuemmler, C., & Bai, C. (2017). Health 4.0: How virtualization and big data are revolutionizing healthcare. In *Health 4.0: How Virtualization and Big Data are Revolutionizing Healthcare* (p. 254). https://doi.org/10.1007/978-3-319-47617-9

Topol, E. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

United Nations, Department of Economic and Social Affairs, Population Division. (2024). *World Population Prospects 2024: Summary of Results*. United Nations. https://population.un.org/wpp/assets/Files/WPP2024_Summary-of-Results.pdf

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

World Health Organization. (2020). *Health workforce* [Web page]. https://www.who.int/health-topics/health-workforce

World Health Organization. (2025). *Artificial intelligence is reshaping health systems: state of readiness across the WHO European Region*. Copenhagen: WHO Regional Office for Europe.

Yu, K. H., Beam, A. L., & Kohane, I. S. (2018). Artificial intelligence in healthcare. *Nature biomedical engineering*, *2*(10), 719–731. https://doi.org/10.1038/s41551-018-0305-z

Chapter 1. **Mind the Gap: Unveiling the Advantages and Challenges of Artificial Intelligence in the Healthcare Ecosystem**

Simona Curiello[1, *], Enrica Iannuzzi[2], Claudio Nigro[2], Dirk Meissner[3]

[1] *Department of Social Sciences, University of Foggia, Italy*

[2] *Department of Economics, University of Foggia, Italy*

[3] *Institute for Statistical Studies and Economics of Knowledge, National Research University Higher School of Economics, Moscow, Russia*

[*] Corresponding Author. Email: simona.curiello@unifg.it

Abstract

Study Design: This study presents a comprehensive systematic literature review (SLR) of the adoption of Clinical Decision Support Systems (CDSSs) in healthcare. We selected English articles dated 2013–2023 from Scopus, Web of Science, and PubMed, found using keywords such as 'Artificial Intelligence', 'Healthcare', and 'CDSS'. A bibliometric analysis was conducted to evaluate literature productivity and its impact on this topic.

Purpose: This work provides an overview of academic articles on the application of Artificial Intelligence (AI) in healthcare. It delves into the innovation process, encompassing a two-stage trajectory of *exploration* and *development* followed by *dissemination* and *adoption*. To illuminate the transition from the first to the second stage, we use Prospect Theory (PT) to offer insights into the effects of risk and uncertainty on individual decision making, which potentially lead to partially irrational choices. The primary objective is to discern whether Clinical Decision Support Systems (CDSSs) can serve as effective means to of "cognitive debiasing", thus countering the perceived risks.

Findings: Of 322 articles, 113 met the eligibility criteria. These pointed to a widespread reluctance among physicians to adopt AI systems, primarily due to trust-related issues. Although our systematic literature review underscores the positive effects of AI in healthcare, it barely addresses the associated risks.

Limitations: This study has certain limitations, including potential concerns regarding generalizability, biases in literature review, and reliance on theoretical frameworks that lack empirical evidence.

Originality/Value: The uniqueness of this study lies in its examination of health care professionals' perceptions of the risks associated with implementing AI systems. Moreover, it addresses liability issues involving a range of stakeholders, including algorithm developers, Internet of Things (IoT) manufacturers, communication systems, and cybersecurity providers.

Keywords: Artificial Intelligence, Healthcare, Clinical Decision Support Systems (CDSSs), Prospect Theory, Risk

Introduction

Innovation – whether technical, technological, or scientific, as outlined by Nonaka (1991) – can be conceptualized in ontological terms as the industrial application of 'inventions'. The delineation between the broader realm of inventions and the more specific category of innovations can be elucidated in terms of the extent of the new concept's diffusion coupled with its commercial and social success. This distinction underscores the transformative journey from conceptualizing novel

ideas (*inventions*) for practical implementation in industry, marking the evolution of innovation and its impact on both economic and societal domains (Schumpeter and Swedberg, 1934).

A broad body of literature has introduced a classification of innovations, drawing a distinction between *incremental* innovations, *radical* innovations, new technological systems, and new techno-economic *paradigms* that are based on their degree of discontinuity with the existing framework (Freeman and Perez, 1986).

As delineated in various manuals addressing innovative processes (Ettlie *et al*., 1984; Dewar and Dutton, 1986), the intensity or degree of innovation can be grouped into three main types: *incremental* innovations that refine existing solutions; *radical* innovations, which replace certain elements with entirely new solutions capable of novel functioning; and *paradigmatic* innovations, which give rise to new economies and fundamentally transform the societal systems within which individuals live, operate, and work.

Within the framework of a conventional firm, radical innovation emerges as a pivotal catalyst for organizational success (McLaughlin *et al*., 2008). This form of innovation enhances a firm's capacity to adapt to a volatile environment and to secure competitive advantages (Tushman and Anderson, 1986). By embracing radical innovation, organizations position themselves to navigate uncertainty, foster adaptability, and gain a strategic edge in the competitive landscape.

In a more complex context, such as the healthcare sector, radical innovation yields multifaceted effects that encompass social, organizational, economic, and technological dimensions. Notably, it can simultaneously or independently exert an effect on other technological innovations, scientific paradigms, organizational behaviour, societal norms, and economic indicators (Caputo *et al*., 2019c; Henderson, 1994; Kelley *et al*., 2013; Tushman and Anderson, 1986; Coccia, 2016; Iannuzzi and Nigro, 2022).

The literature has consistently classified the innovation process into two stages: the initial step involves innovation *development* while the second includes *dissemination*, *adoption*, *implementation*, *maintenance*, *sustainability*, and *institutionalization* (Orlandi *et al*., 1990; Oldenburg and Parcel, 2008).

Organizations first develop and refine innovations during the *development* stage. In the second stage, *dissemination* focuses on convincing specific target groups to adopt innovations. *Adoption* considers the requirements and beliefs of potential adopters, as well as their responses to innovation and factors that might make them more or less likely to adopt it. Strategies that encourage behavioural changes and overcome obstacles are crucial. The decision to adopt a certain innovation can be affected by 1) *awareness* of it, 2) *procedural knowledge* of its use, and 3) *conceptual knowledge* of how it functions (Takahashi and Vandenbrink, 2004; Rogers *et al*., 2014).

Proceeding to the research main topic, Artificial Intelligence (AI) has emerged as a transformative technology with the potential to revolutionize various industries, including healthcare (Johnson *et al*., 2014).

Through the utilization of Machine Learning (ML) algorithms, deep neural networks, and data analytics, Artificial Intelligence (AI) offers unprecedented capabilities for processing and interpreting vast volumes of medical information (Caputo *et al*., 2023a). This can enhance clinical decision-making, improve patient outcomes, and foster more efficient healthcare systems (Esteva *et al*., 2019; Topol, 2019).

At present, the integration of AI into healthcare organizations is in the initial stage of adoption, the *development* (or exploration) stage (Del Giudice *et al*., 2021; Ulloa *et al*., 2022). Its widespread implementation by hospitals and other health facilities has not yet materialized, possibly because of the range of challenges intricately associated with these innovative systems, as articulated by Rogers and colleagues (2021). Notably, there is a pervasive lack of awareness, procedural knowledge, and conceptual knowledge among healthcare practitioners, organizations, and institutional actors regarding the adoption of this innovative technology.

To better comprehend the issues and limitations of this implementation, it is essential to point out that, at this developmental juncture, the ultimate responsibility for any adverse events stemming from the adoption of AI systems rests with the decision maker, be it the medical professional or healthcare facility. Consequently, the need for regulations and certifications encompassing technological and operational standards becomes paramount. Such measures are crucial for delineating the risks and responsibilities associated with AI systems' deployment. This point holds a central position on the agenda of governments and supra-governmental organizations, with the European Union playing a leading role in addressing and shaping the regulatory landscape.

What are the areas that academics should concentrate on when considering the possibility of complete *substitution* of human skills by increasingly advanced machines (Del Giudice *et al*., 2023)? When implementing complex AI systems in healthcare, it becomes crucial to contemplate the need for a *risk-reward assessment*. How is the *cost-benefit* profile managed when adopting AI which, in theory, could replace humans in clinical diagnosis?

To date, scientific research has tended to focus on abstract comparative assessments, such as the "*horse and train*" metaphor. However, it is crucial to identify the source of *risk*, especially when applying these concepts in a clinical setting.

When investigating spillovers in healthcare, researchers must consider aspects related to the adoption and development of innovative systems, including the complexity of AI systems, which consist of a combination of *facilities* such as 1) algorithms, 2) devices (IoT), 3) communication

systems, and 4) services. Furthermore, implementing embedded AI systems affects organizational processes and performance in the healthcare sector.

The objective of this research is to map the intellectual landscape of AI in healthcare. This will be achieved by identifying key contributors, prevalent topics, and the extent of cross-disciplinary collaboration. This study aims to examine the way scholars integrate AI systems, considering components such as algorithms, the Internet of Things (IoT), communication systems, and services. Furthermore, the study analyzes the literature that inspects the advantages and disadvantages of AI adoption from the perspectives of both physicians and patients. Indeed, a deeper emphasis on understanding the *risks* associated with these systems is necessary, particularly in relation to the uncertainty of the benefits for end users (healthcare professionals and patients). In light of the considerable risks and uncertainty inherent in the field of healthcare, any configuration error potentially has a significant impact on patients' health, particularly in the areas of diagnosis, treatment, and prevention. Healthcare professionals are exposed to dual risks: 1) *adverse patient outcomes* and 2) *legal consequences* resulting from their decisions.

Kahneman and Tversky's Prospect Theory (1979) provides a well-established framework for the analysis of decisions under conditions of risk and uncertainty. The theory elucidates the manner in which decision makers assess outcomes in relation to a *reference point*, perceiving them as gains or losses. This is particularly pertinent in clinical settings, where decisions regarding the adoption of novel technologies, such as CDSSs, necessitate weighing the risk of an AI system providing an inaccurate diagnosis or treatment recommendation against the potential benefits of enhanced outcomes. Moreover, this study employs Prospect Theory to investigate the potential influence of cognitive biases on healthcare professionals' risk appetite when adopting novel AI systems in high-stake contexts (Gisbert-Pérez *et al*., 2022).

With regard to the above considerations and starting from a preliminary reconstruction of the theoretical background related to Prospect Theory (PT), this work is intended to present the findings of a systematic literature review through the lens of a theoretical framework dealing with cognitive psychology applied to risky decisions in AI adoption in healthcare processes. The theoretical framework of PT takes over in the *adoption* and *dissemination* phases. In fact, in the logic of *exploitation*, Prospect Theory becomes the cornerstone for risk-opportunity assessment, since human decision-makers need to face liability if harm is caused to a patient, *as long as no mandatory regulatory framework is adopted*.

This study attempts to contribute to the field through comprehensive bibliometric analysis by identifying key gaps in existing literature. In particular, this paper highlights the dearth of discourse on the ramifications of AI integration in clinical decision-making on healthcare structures from an

organizational standpoint and on clinicians from a professional perspective. Prospect Theory offers valuable insights into the psychological factors influencing AI adoption in this context. This research contributes to the existing body of knowledge by offering a comprehensive analysis of the intersection between AI and healthcare. Although previous studies have examined the technical and medical aspects of AI adoption, this study makes a unique contribution by applying Prospect Theory to investigate the psychological and behavioural factors influencing decision making among healthcare professionals. By focusing on the cognitive biases that arise when clinicians are faced with adopting AI systems under conditions of risk and uncertainty, this study illuminates the previously underexplored dimensions of AI integration. Furthermore, this study presents a novel framework for analyzing the influence of risk perception, loss aversion, and framing effects on healthcare professionals' willingness to integrate AI into their clinical workflows (Westerbeek *et al*., 2021; Choudhury, 2022). This research indicates the possibility of AI systems functioning as instruments for cognitive debiasing, thereby facilitating more rational decision-making processes (Bennett and Hauser, 2013). In doing so, it not only addresses a gap in the existing literature, but also provides valuable insights for healthcare organizations, policymakers, and AI developers, which can inform the development of more effective and responsible AI adoption in healthcare settings (Khanijahani *et al*., 2022).

Theoretical Framework

The paradigmatic change driven by AI

Here, we draw a parallel, likening the contemporary wave of AI-driven innovation to the emblematic 1830 race between a steam-powered locomotive and a horse-drawn carriage, a historic event that epitomized the transformative era of the Industrial Revolution. The race took place on August 28, 1830, between Peter Cooper's small Tom Thumb locomotive and the horse-drawn Baltimore and Ohio (B&O) Railroad car. The locomotive showcased the advantages of steam power, reaching 15 miles per hour. This metaphor illustrates the ongoing comparison between human and AI capabilities, reminiscent of the historic race's focus on the speeds of transportation modes. Scholars often analyze the reliability and alignment between doctors' judgments and AI suggestions, emphasizing that the initial phase of AI development has concentrated on comparing human capabilities to those of machines (Mahadevaiah *et al*., 2020; Franklin, 2012; Saviano *et al*., 2023).

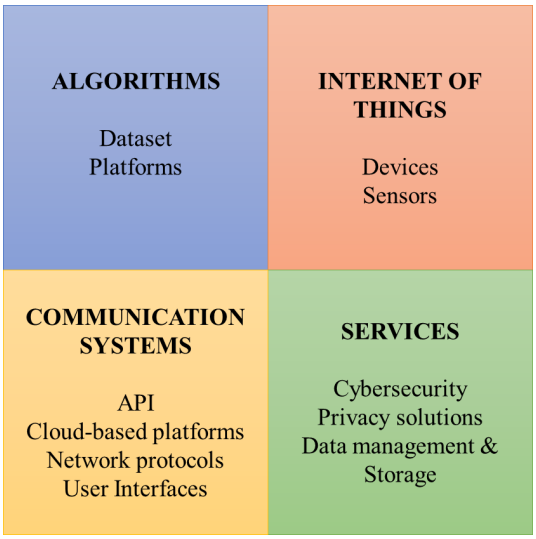
To comprehend why AI cannot be approached in a process-innovation logic (especially in the social and health context), it is helpful to consider the radical perspective of certain philosophers and specialists. During a debate with Yann Le Cun, Chief AI Scientist at Meta and one of the founding fathers of AI, the bestselling author Yuval Noah Harari stated: "*It's not a knife: a knife can't decide*

whether to kill someone or to save life in surgery. The decision is always in human hands. AI is the first tool that potentially can replace us in decision making"[2].

AI displays supremacy over people in an ever-growing range of capabilities such as detecting human emotions. It is crucial to recognize that the AI revolution is not just a consequence of faster, more sophisticated computers, but of advancements in the biological and social sciences. Over the next century, the convergence of biotechnology and AI could trigger the emergence of physical, mental, and bodily characteristics that differ greatly from their usual human form (Harari, 2019). Experts such as Kurzweil (2005), Bostrom (2014), Brynjolfsson and McAfee (2014), Tegmark (2017), and Russell and Norvig (2021) have emphasized how the emergence of AI systems may represent a paradigm shift, with the potential to outperform human decision-makers in tasks such as analysis and intervention.

Scherer (2015) highlighted four crucial domains for AI integration: *algorithms*, *IoT*, *communication systems*, and *services*. The management of AI systems requires comprehensive consideration of these components, which redefine tasks that have traditionally required human intelligence.

Figure 2 – AI Integrated System



Source: Authors own work

Algorithms serve as the cornerstone of AI systems and involve two key parts: dataset and platform. An algorithm provides the step-by-step instructions needed for computers to execute specific tasks and make specific decisions, process, and analyse data to extract patterns and make predictions.

[2] Yuval Noah Harari (Sapiens) versus Yann Le Cun (Meta) on Artificial Intelligence (2023) – https://www.lepoint.fr/sciences-nature/yuval-harari-sapiens-versus-yann-le-cun-meta-on-artificial-intelligence-11-05-2023-2519782_1924.php

The *Internet of Things* (IoT) is a network of interconnected devices and sensors that exchange data (Kishor and Chakraborty, 2022). In AI systems, the IoT plays a central role, providing real-world data that can be used for analysis and informed decision-making.

Within AI systems, there are several basic *IoT components*, including:

- *Sensors*: physical devices that collect data from the environment;
- *Edge computing*: processing data closer to the source (i.e., in the device or sensor) rather than sending it to a central server;
- *Data connectivity*: IoT devices typically use various communication protocols (e.g., Wi-Fi, Bluetooth, and cellular) to transmit data to central servers or other devices for further processing;
- *Data storage*: databases or cloud platforms, making it accessible for AI algorithms to analyse and derive insights.

Communication Systems constitute the third facility that enables data exchange between AI models, IoT devices, and external entities. The key components of *communication systems* are as follows:

- *Application programming interfaces* (APIs), facilitating software component communication;
- *Cloud services*, ensuring infrastructure for AI models;
- *Networking protocols*, enabling secure data transmission;
- *User interfaces* for human interaction.

Services encompass cybersecurity, data management, and privacy. For instance, the ENISA actively addresses AI cybersecurity challenges by offering security recommendations.

These components synergize to create functional AI systems that can analyse IoT data, make intelligent decisions using algorithms, and effectively communicate the results or actions. This framework minimizes risks and errors, particularly in fields such as biomedicine.

Table 1 - Facilities constituting AI systems

Facility	Actors Involved	Main Responsibility	Role
Algorithms	Data Scientists and Machine Learning Engineers	Developing and training AI algorithms	Focus on the creation and optimization of machine learning models and data analysis.
	Software Developers and Engineers	Integrating AI algorithms into software applications	Develop the software components that execute and interact with AI algorithms
IoT	IoT Engineers and Hardware Specialists	Designing and maintaining physical devices and sensors	Focus on the hardware aspect of IoT, including sensor selection and integration
	Data Engineers	Managing and processing data from IoT devices	Design data pipelines, databases, and ensure data quality for IoT data.
Communication Systems	Software Developers and Engineers	Creating user interfaces and integrating AI with communication systems	Develop APIs and user interfaces for data exchange and system

			communication
	Cloud Architects and DevOps Engineers	Managing infrastructure and deployment, including communication systems	Ensure that data and AI algorithms are deployed effectively and securely in the cloud
	Project Managers and Product Managers	Overseeing the entire development and implementation process, including communication aspects	Ensure that the project aligns with business goals and is delivered on time and within budget
	Quality Assurance (QA) Testers	Testing the AI system, including communication functionalities	Validate system performance and functionality to identify and address issues
	Business and Legal Experts	Addressing business and legal considerations related to data and communication	Ensure compliance with legal requirements, data privacy, and business strategy
	End Users	Interacting with and benefiting from the AI system	Provide feedback and usage data, which can be used to improve the system's communication and user experience
Services	Data Engineer	Design and maintain data storage systems; create and manage data pipelines; ensure data quality and availability; transform and integrate data	Data engineering and management
	Database Administrator	Manage and maintain database systems; optimize database performance; perform backup and recovery operations; ensure data security	Database administration and management
	Data Analyst	Extract insights from data; analyse and report on data; visualize data for decision-making; collaborate with teams to leverage data effectively	Data analysis and reporting expert
	Data Scientist	Develop machine learning models; extract patterns and insights from data; explore data for actionable insights; hypothesis testing and predictive modelling	Data science and machine learning
	Chief Information Security Officer (CISO)	Oversee data security strategies; ensure compliance with data protection regulations; develop and implement cybersecurity policies and practices	Leadership role in cybersecurity and data protection
	Cybersecurity Analyst	Monitor and analyse security threats; implement security measures and protocols; investigate and respond to security incidents	Cybersecurity and threat analysis
	Data Privacy Officer	Ensure compliance with data privacy regulations (e.g., GDPR); handle data subject requests; develop and enforce data	Specialist in data privacy and compliance

		protection policies and practices	

Source: Authors own work

Adoption of AI in healthcare

The integration of AI into healthcare, specifically in clinical decision support, has significantly affected disease diagnosis, treatment prediction, and the identification of complications (Middleton *et al*., 2016; Lundberg *et al*., 2018). Regulatory frameworks have been established to govern the use of Clinical Decision Support Systems (CDSSs) in healthcare globally. Notable examples include risk-based categories for Software as a Medical Device (SaMD) established by the US FDA, which ensure premarket review, post-market surveillance, and quality system regulation (US FDA, 2018). Similarly, the GDPR mandates strict data protection principles for AI systems in healthcare (European Parliament and Council of the European Union, 2016). In addition, the Ethics Guidelines for Trustworthy AI of the European Commission (European Commission, 2019) emphasize transparency and accountability.

The EU AI Act (2023) is a globally pioneering comprehensive body of legislation that addresses potential risks to health, safety, fundamental rights, and environmental concerns while emphasizing democracy and the rule of law. Its field ranges from medical imaging analysis, which might assist radiologists in detecting abnormalities, to Natural Language Processing (NLP) techniques, which aim to extract insights from extensive medical data (Esteva *et al*., 2019; Rajkomar *et al*., 2019).

Although AI demonstrates great potential, many ethical and technological concerns still need to be addressed (Siala and Wang, 2022; Camilleri, 2024). It is crucial to attend to *privacy*, *consent*, *bias*, and *accountability* if trust is to be maintained and medical practice standards are upheld (Babushkina and Votsis, 2022; McLennan *et al*., 2022; Smallman, 2022; Carter *et al*., 2020). Technical challenges in AI systems include *reliability*, *interpretability*, *transparency*, and the need for accurate *input data* to avoid bias (Obermeyer *et al*., 2019; Xu *et al*., 2023; Valente *et al*., 2022).

While CDSSs aim to 'cognitively debias' clinicians, their human-machine interaction introduces risks. In particular, experts' reluctance to adopt them on account of their high-risk nature may be explained by the following factors:

- the system's functioning itself, because many ML systems are black-box systems that deliver a result or a decision without explaining or showing how they did so (McCoy *et al*., 2022). This is potentially risky when making vital decisions for patients;
- the liability problem arises because of the involvement of multiple actors in the development and deployment of systems equipped with AI (e.g., software development companies, manufacturers, and healthcare providers). When a patient is harmed as a result of a malfunctioning or hacked

system, it is likely to be difficult to immediately identify a responsible actor or even a reasonable way of sharing risk among the various parties;

- the ethical and legal issues related to the adoption of these systems impose a need for cybersecurity systems on healthcare structures that deal with vast amounts of patient data (privacy and data protection, data breaches) (Ramgopal *et al*., 2023; Prakash *et al*., 2022);
- a relative lack of standards that certify the reliability of systems gives rise to trust issues concerning the system's recommendations and a regulatory gap regarding AI applications as decision-makers in healthcare (Pee *et al*., 2019; Raja and Asghar, 2020; Senbekov et al., 2020; Choudhury *et al*., 2022).

An emerging challenge in AI adoption in healthcare is the need for a regulatory framework. In the event of large-scale adoption of AI systems, European intervention may be necessary, particularly in terms of:

- definition of standards (Hawkins *et al*., 2017);
- introduction of standard certification mechanisms (Heesen *et al*., 2020);
- risk-assessment and risk-coverage mechanisms (Tabassi, 2023).

Innovation, particularly in high-stakes fields such as healthcare, is inherently prone to risk. Failure is often an unavoidable consequence of innovative projects (Maslach, 2015; D'Este *et al*., 2016; Jensen *et al*., 2016; Sreen *et al*., 2024). Some scholars posit that the probability of failure increases in accordance with the intensity and radical nature of innovation (Kamoto, 2017; Sharma *et al*., 2017). In the case of AI systems, such as CDSSs, the potential for innovation failure is a critical factor to consider, given the complex and uncertain environments in which these systems operate. Failures in AI-driven healthcare innovations such as inaccurate diagnoses or treatment suggestions can have significant implications for patient outcomes and organizational liability. However, as Carroll and Mui (2008) observed, failure can also be a powerful source of learning and progress. Indeed, nearly half of all failures offer underexploited opportunities for technological advancement and social welfare. In this regard, *backward-looking strategies* that learn from these failures can significantly enhance *forward-looking innovation strategi*es, thereby leading to a more robust and effective implementation of AI in clinical practice (Rhaiem and Amara, 2021; Alowais *et al*., 2023; Houfani *et al*., 2021).

Prospect Theory in clinical decision-making

Prospect Theory (PT), proposed by Kahneman and Tversky, provides insights into the decision-making processes. It points out that individuals tend to be risk-averse in gain situations and risk-seeking in loss scenarios. According to PT, clinical decision-making involves two different types of

mental processes: intuitive (System 1) and analytical (System 2), which are affected by the complexity of a given situation (Kahneman and Tversky, 1979).

Croskerry's diagnostic reasoning model helps to understand how clinicians collect, process, and interpret patient information (Kovacs and Croskerry, 1999). Medical decisions are frequently made in uncertain and dynamic environments, which can lead to errors. Decisions that have the potential to harm patients are considered *high-risk* because they can result in diagnostic failures by clinicians (Brennan *et al.*, 1991).

Although there is little prospect that these concepts can be applied immediately in healthcare contexts and medical decisions, it is possible to extend the notion of risk as "uncertain outcome" to that of risk as "severe consequences" deriving from certain behaviours (Rothman and Salovey, 1997). When adopting CDSSs, there is a complex relationship between *risk-perception* and AI utilization. The absence of a regulatory framework amplifies the psychological impact of risk due to the adoption of AI, which highlights the need for cognitive mapping to navigate emerging issues in the utilization of AI technology.

The adoption of such systems could constitute an approach also known as "debiasing strategy", in the sense of switching from pattern recognition to analytic thinking triggered by a stimulus (Daniel *et al.*, 2018; Harada *et al*., 2020).

The basic idea behind this contribution is that, even if it is true that the perception of uncertainty with respect to the patient's health status (*loss-frame*) could result in a greater propensity to risk and, consequently, a reliance on AI algorithmic systems, "severe consequences" may occur on the doctor at the same time, taking into account legal issues arising from potential damage to the patient. Thus, the incentive to adopt such systems could result from the mitigation of risk, understood as the legal liability of doctors.

The PT framework is particularly useful in this first phase of AI research and development. In fact, in the absence of large-scale diffusion and adoption of these technologies, the lack of a regulatory framework places the psychological impact of risk in the hands of a person potentially interested in adopting AI systems. A study of the contributions to the subject makes it possible to define a cognitive map capable of highlighting these aspects:

- the areas of application of AI technologies in this first stage of *exploration* and *development*;
- the issues that emerge in the *exploitation* and *diffusion* of AI technologies.

From this perspective, the aim is to determine what is known about:

- macro areas of risk: *high-risk* vs. *low-risk* decision making;
- the facilities that qualify an integrated AI system.

Knowledge Management and Change Management

In light of the pivotal role played by *dynamic capabilities* – which refer to an organization's ability to integrate, build, and reconfigure internal and external competencies to address rapidly changing environments (Korherr and Kanbach, 2023; Mele *et al*., 2024; Capolupo *et al*., 2023) – and Big Data Analytics (BDA) – a set of advanced techniques and tools that allow organizations to process vast volumes of complex data for more informed decision-making (Felicetti *et al*., 2024) – it is imperative to integrate the tenets of *Knowledge Management* (KM) and *Change Management* (Truong *et al*., 2024; Kettinger and Grover, 1995).

Dynamic capabilities facilitate firms' ability to remain agile and adaptable, particularly in sectors such as healthcare, where new technologies are continuously evolving (El Morr and Subercaze, 2010; Mele *et al*., 2023). The concept of dynamic capabilities, as developed in the field of KM, examines the ways in which organisations can leverage both tacit and explicit knowledge to gain competitive advantages, particularly in the context of complex systems such as business decision-making algorithms (BDA) and AI (Teece, 2023). Similarly, BDA enables organizations to derive actionable insights from extensive datasets, thereby facilitating innovations in clinical decision-making, operational efficiency, and patient care. Nevertheless, the effective implementation of these tools hinges on the establishment of robust KM systems that guarantee the capture, dissemination, and utilization of knowledge throughout the organization. Additionally, *Change Management* practices are essential for ensuring the seamless integration of these novel technologies into existing workflows and for overcoming resistance to change (By, 2005). In the context of health care, KM emphasizes the role of knowledge acquisition, storage, sharing, and utilization in fostering innovation and improving decision-making processes (Khan *et al*., 2023; Magni *et al*., 2023a). In conjunction with Change Management theories, these frameworks can assist in elucidating the manner in which healthcare organizations can integrate AI-driven innovations while maintaining adaptability.

Moreover, a significant number of healthcare organizations continue to grapple with the effective management of knowledge as employees frequently exhibit reticence to disseminate information or expertise among their colleagues. The concept of *knowledge hiding* was initially introduced by Connelly *et al*. (2012), who defined it as a deliberate refusal to share knowledge with others despite requests for assistance (Wilson *et al*., 2021). Such requests could relate to guidance on tasks, professional insights, or practical skills (Caputo *et al*., 2021). In the context of healthcare, the advent of AI has introduced an additional layer of complexity. This behaviour is a strategic attempt by employees to sabotage AI processes, thus mitigating the risk of job displacement. This underscores the importance of addressing the psychological and social dimensions of AI adoption in healthcare, ensuring that clinicians feel secure and supported in their roles.

As posited by Arias-Pérez and Vélez-Jaramillo (2022), the accelerated integration of AI can precipitate a state of technological turbulence, thereby engendering employees' awareness of potential threats to their job security. This can prompt various forms of knowledge hiding, whereby employees withhold critical information to protect their roles or mitigate perceived risks associated with job displacement (Arain *et al*., 2022; Al Hawamdeh, 2023; Bhatti *et al*., 2023; Chhabra and Pandey, 2023; Wu *et al*., 2023; Liao *et al*., 2024).

The withholding of knowledge in this manner represents a strategic action that may undermine the functionality of AI systems, thereby protecting employees from potential job redundancies (Caputo *et al*., 2021). Historically, human resources training has been oriented towards a competitive perspective, wherein individual success is contingent upon outperforming colleagues or controlling access to essential resources for effective work processes (Long *et al*., 2024). Consequently, withholding knowledge can become a deliberate strategy, whereby knowledge is regarded as a valuable internal asset. Nevertheless, within the context of healthcare, such tendencies towards the concealment of knowledge may prove to be an obstacle to the successful implementation of AI-driven initiatives that rely on the free flow of information (van de Wetering and Versendaal, 2018).

The integration of KM reinforces the argument for the potential of AI by aligning the adoption of technology with knowledge-intensive processes (Shahmoradi *et al*., 2017). AI systems that support CDSSs depend on the utilization of extensive medical data, which directly correlates with the KM principles of knowledge codification and transfer. As healthcare institutions adopt AI, they undergo transformations that affect both knowledge flow and organizational routines. This transition requires the integration of Change Management practices to mitigate resistance and enhance acceptance among the staff. Scholars have emphasized that when implementing disruptive technologies such as AI, the human factor of resistance to change can be mitigated by KM systems that support continuous learning and adaptation to new processes.

The integration of AI-driven systems in the healthcare sector is consistent with the core objective of KM, namely the utilization of institutional knowledge to optimize outcomes, minimize errors, and enhance patient care. It has been demonstrated that the successful implementation of AI is contingent upon an organization's capacity to manage knowledge in an efficacious manner, including through the utilization of continuous learning and feedback loops.

Research Questions and Hypothesis

By synthesizing theoretical insights into the transformative potential of AI in healthcare and the paradigm shifts that this introduces, a set of streamlined research questions (RQs) and hypotheses (Hps) have been developed to bridge theory with empirical investigation. We refer to the research

design proposed by Zupic and Čater (2015), which presents a number of research questions that can be explored through a bibliometric study. Such a study concentrates on authors, journals, keywords, and citations.

Considering the first stage of innovation (*exploration* and *development*), the following questions aim to map the intellectual terrain of AI in healthcare, identifying leading voices (RQ1a), prevailing topics (RQ1b), and the extent of cross-disciplinary engagement (RQ1c), reflecting AI's pervasive influence across various domains.

RQ1a. *Who are the most prominent authors in the research field*? (citation analysis)

RQ1b. *What are the main topics related to the application of AI in healthcare on which the literature focuses*? (co-word analysis)

Moreover, as suggested by Javaid *et al*. (2019), the effects of AI on the medical domain are substantial and require proficiency in various scientific fields. It would thus be advantageous to explore the question of (**RQ1c**) *whether authors cooperate across research domains during the preliminary phase of innovation* (co-author analysis).

In the second stage of innovation (exploitation and diffusion/adoption):

RQ2. *Do scholars of this discipline look at the phenomenon jointly considering the four facilities (Algorithms, IoT, Communication systems and Services)*?

Through the formulation of these RQs, the corresponding research hypotheses can be summarized as follows:

Hp1. The literature predominantly views AI as a technical tool, with limited discussion of its broader methodological impact on healthcare practices.

Hp2. Authors on this topic focus on AI systems, jointly considering all structures as a bundle.

HP3. Adopting the theoretical framework of Prospect Theory:

HP3a. Academics highlight the benefits and performance improvements of adopting AI in the healthcare sector, leaving aside the potential concerns and harms of these systems.

Hp3b. Academics focus only on the risk perception of AI adoption in healthcare from the perspective of physicians and patients.

Hp4. Researchers deal with AI by focusing on an organizational level, starting from the assumption that structural factors such as organizational *size*, *workflows*, *training* and *security* could play a key role in the adoption of AI technologies.

Methodology

A systematic literature review (SLR) is guided by inductive reasoning that involves establishing and applying a specific set of criteria to construct a document corpus that scholars can evaluate. We

employ this methodology to examine the current literature in this field (Ammirato *et al*., 2022). Following Kraus *et al*. (2022), this review can be classified as a "between-domain hybrid", because it deals with a certain topic (*Artificial Intelligence*) in a specific field (*healthcare*).

In particular, our approach consists of a systematic literature review based on the methodological steps outlined by Massaro *et al*. (2016), Choudhury and Asan (2020), and Ali *et al*. (2023).

We adopted the SLR methodology because it allows the integration of both qualitative and quantitative data analysis methods. Combining SLR with bibliometric analysis, which involves quantifying and evaluating literature through algorithms and statistical methods, provides an objective means of assessing large collections of articles. This method allows for the examination, classification, and investigation of large datasets, often consisting of hundreds or more members. We describe a step-by-step approach for conducting the review, which involves defining research questions, creating review protocols, screening and selecting articles based on predetermined criteria (Dabić *et al*., 2021), extracting relevant data, and synthesizing findings (Donthu *et al*., 2021).

In line with the scoping review methodology, the inclusion and exclusion criteria were established at the outset and refined during the iterative screening process to ensure consistency throughout the research.

To guarantee the relevance and quality of the selected studies, the following inclusion criteria were established (Kitchenham *et al*., 2009):

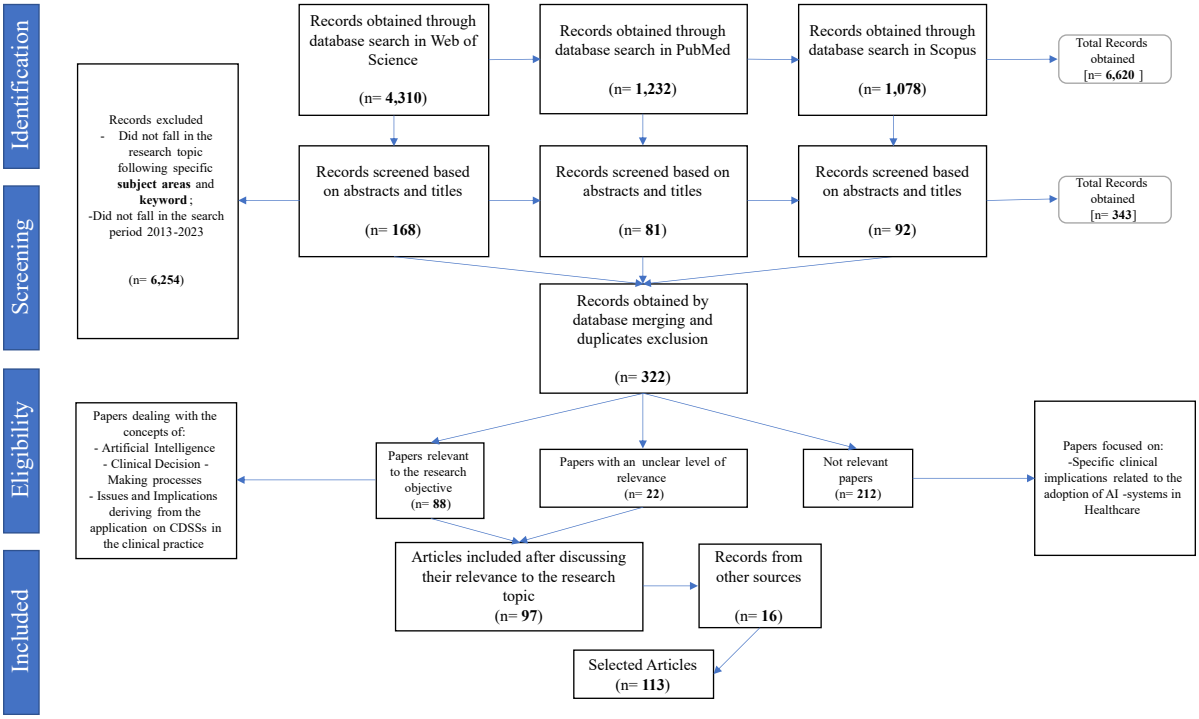
1. *Research Focus*. Articles must specifically address the integration or application of Artificial Intelligence (AI) in healthcare, focusing on Clinical Decision Support Systems (CDSSs).
2. *Publication Date*. Only articles published within the last decade (2013–2023) were considered in order to capture the most recent developments and trends.
3. *Language*. Articles must be written in English to ensure accessibility and comprehensibility.
4. *Article Type*. Peer-reviewed journal articles were prioritized to ensure the inclusion of high-quality and validated studies. Grey literature, conference proceedings, books, and book chapters were excluded.
5. *Database Selection.* Articles were selected from recognized and reputable databases, such as Scopus, Web of Science, and PubMed. This selection was based on comprehensive coverage, reliability, and relevance to the research focus. PubMed has rapidly become synonymous with medical literature searches worldwide. Its extensive collection of biomedical literature, including articles from MEDLINE, life science journals, and online books, makes it an indispensable resource for healthcare research (Falagas *et al*., 2008); Scopus covers a wide range of topics, including physical sciences, health sciences, life sciences, and social sciences (Falagas *et al*., 2008). Furthermore, it was chosen based on limitations, such as those of Guo

et al. (2020), who suggested that future studies should include databases such as Scopus to explore more potential papers and expand the scope of their research. Finally, the inclusion of Web of Science ensures access to a wide range of influential and peer-reviewed articles because of its robust citation indexing and analysis tools, which appear to be crucial for a thorough understanding of the adoption and impact of AI and CDSSs in healthcare (Bakkalbasi *et al*., 2006).

Exclusion criteria were devised to filter out studies that did not meet the requisite standards or relevance for this review. Studies that focused solely on the technical or medical implications of AI without specifically addressing the core issues of AI adoption in healthcare and its impact on healthcare structures and professionals were excluded.

According to these inclusion criteria, the study relied on articles from scientific journals, excluding grey literature, conference proceedings, books, and book chapters. The search, conducted on May 17, 2023, included a language filter and excluded articles not in English. Publication date limits were set for papers published before 2013 across databases, providing a review of the last decade. The research addresses three main questions: identifying key authors and journals through citation analysis (Zhao and Strotmann, 2008; White and McCain, 1998; White and Griffith, 1981; Gao and Guan, 2009; Small and Koenig, 1977; Yan and Ding, 2012; McCain, 1991), exploring the state-of-the-art and major applications of AI in healthcare, and assessing cross-contamination among seemingly distant research fields. The second step involved developing a research protocol based on the existing knowledge and literature, narrowing the search to papers aligned with specific features, and following the PRISMA workflow (Tricco *et al*., 2018) for systematic literature reviews (Raymond *et al*., 2022; Srivani *et al*., 2023).

Figure 3 - Preferred Reporting Items for Systematic Review and Meta-Analysis (PRISMA) flow diagram. This PRISMA flowchart shows the number of records collected at each stage of the exploratory literature review

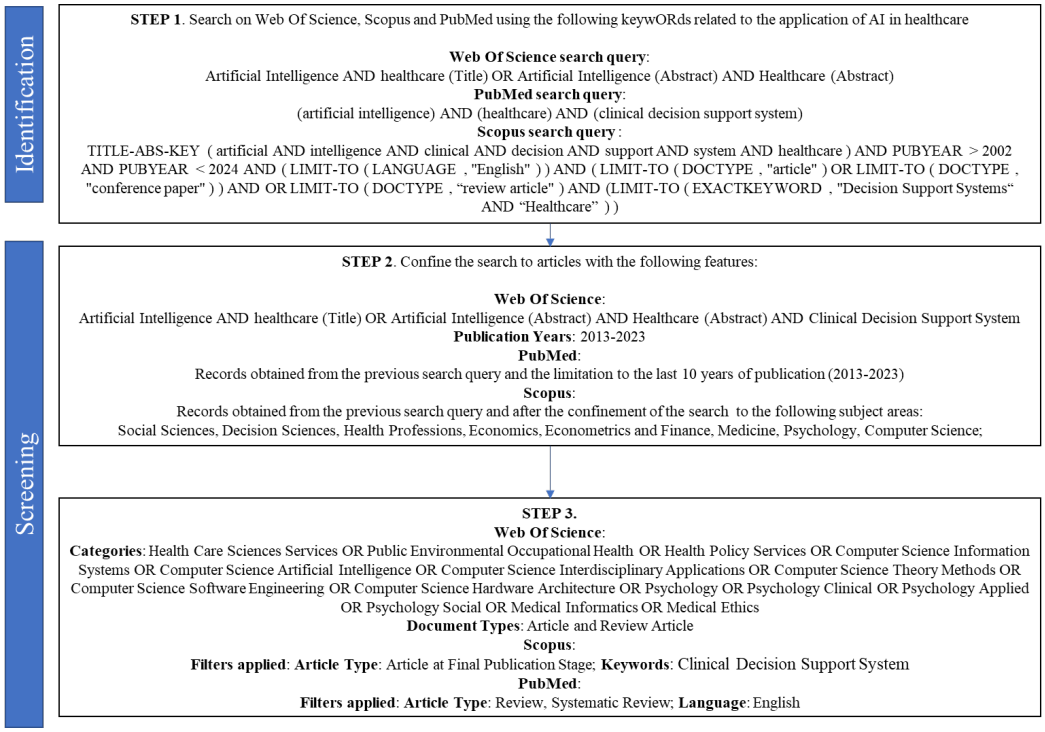


Source: Authors own work

In the second step, inclusion and exclusion criteria were defined and the search strategy was specified. According to a study by Cobo *et al.* (2011), numerous online bibliographic databases store metadata related to scientific works and can serve as valuable sources of bibliographic information, such as Clarivate Analytics Web of Science (WoS), accessible at http://www.webofknowledge.com, Scopus at http://www.scopus.com, Google Scholar at http://scholar.google.com, and PubMed, available at https://pubmed.ncbi.nlm.nih.gov/. (Cobo *et al.* 2011). In particular, the PubMed and WoS databases were chosen following the suggestions of Secinaro *et al.* (2021), who recommended that "*further analysis could also investigate the PubMed, IEEE, and Web of Science databases*".

The third step involved screening and selection, guided by a *top-down* approach to keyword identification, initially yielding 6620 articles. Subsequent refinement narrowed the search to meet specific criteria and a defined period, resulting in a final dataset of 322 articles. The relevance criteria focused on papers addressing AI in healthcare with a particular emphasis on CDSSs and key topics, such as *ethical and legal issues*, *impact on medical workflow*, and *clinicians' attitudes*. Of the 322 articles, 88 were considered highly relevant, 22 initially had unclear relevance, and 212 were excluded. Further analysis of the unclear relevance category led to the inclusion of nine additional articles. Additionally, 16 records from other sources (Google Scholar) were incorporated, bringing the final count to 113 papers selected for the review.

Figure 4 - Screening Phase



Source: Authors own work

Results

Bibliometric Analysis

Bibliometric analysis has five stages: study design, data collection, analysis, visualization, and interpretation (Zupic and Čater, 2015).

The fourth step of our research process involved extracting data from the dataset. The following table 2 contains a summary of the main information retrieved from the bibliographic inputs included in the *Bibliometrix* software. All bibliometric networks can be graphically visualized or modelled, and the *networkPlot* function plots a network created by *biblioNetwork* using R routines (Zhao and Strotmann, 2008; White and McCain, 1998; White and Griffith, 1981; Gao and Guan, 2009; Small and Koenig, 1977; Yan and Ding, 2012; McCain, 1991).

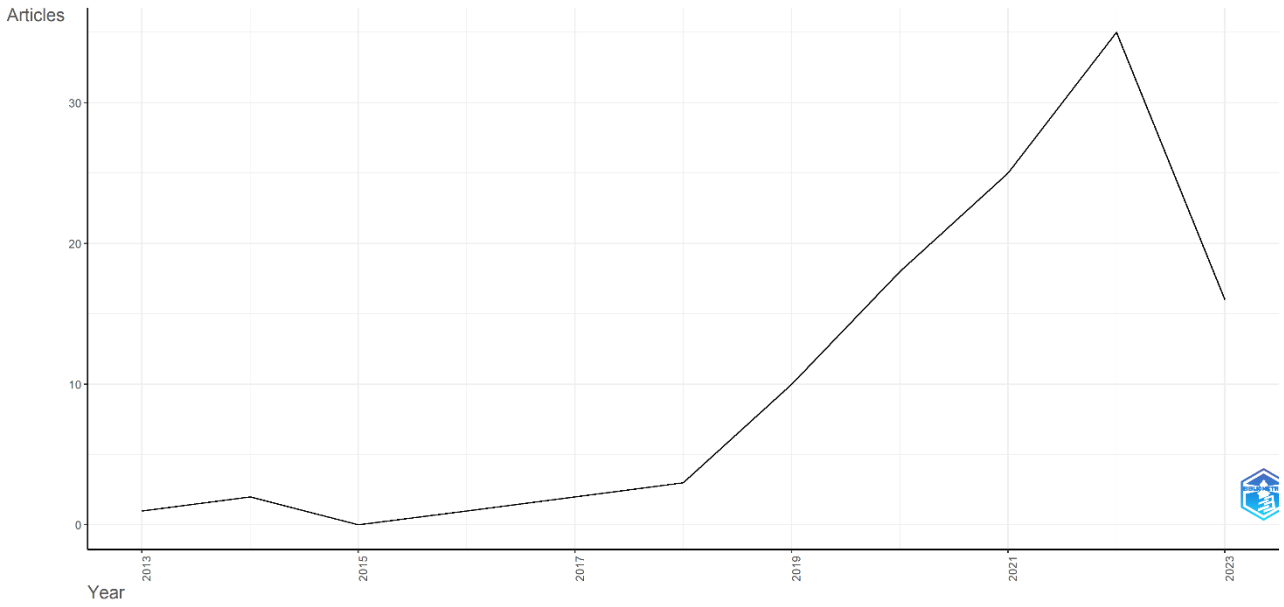
The analysis spanned a decade (2013–2023), with an annual growth of 31.95% in the output of articles. Production increased notably after 2018 (Figure 4). On average, each article was written by approximately five authors with a collaboration index (CI) of 4.96. This synthesis highlights key findings from the analysis, offering insights into trends, collaboration patterns, and growth in research over the study period.

Table 2 - Main Information from database

Description	Results
Main Information about Data	
Timespan	2013-23
Sources (Journals, Books, etc)	74
Documents	113
Annual Growth Rate %	31,95
Document Average Age	2,13
Average citations per doc	49,79
References	7691
Document Contents	
Keywords Plus (ID)	882
Author's Keywords (DE)	360
Authors	
Authors	517
Authors of single-authored docs	11
Authors Collaboration	
Single-authored docs	11
Co-Authors per Doc	4,89
International co-authorships %	38,05
Document Types	
article	69
book chapter	1
conference paper	2
note	2
review	38
short survey	1

Source: Authors own work

Figure 5 - Annual Production of Scientific papers in the research area



Source: Authors own work

Table 3 summarizes the local impact of the top ten journals and sources included in the search sample, focusing on the h-index, the total global impact of each of them (times cited, TC), and the number of papers (Hirsch, 2005). These indexes refer to a *performance* analysis that enables the authors of the present work to assess productivity in terms of publication and the impact, as measured by the number of citations, of the literature relating the domain through these quantitative metrics (Kraus *et al*., 2022).

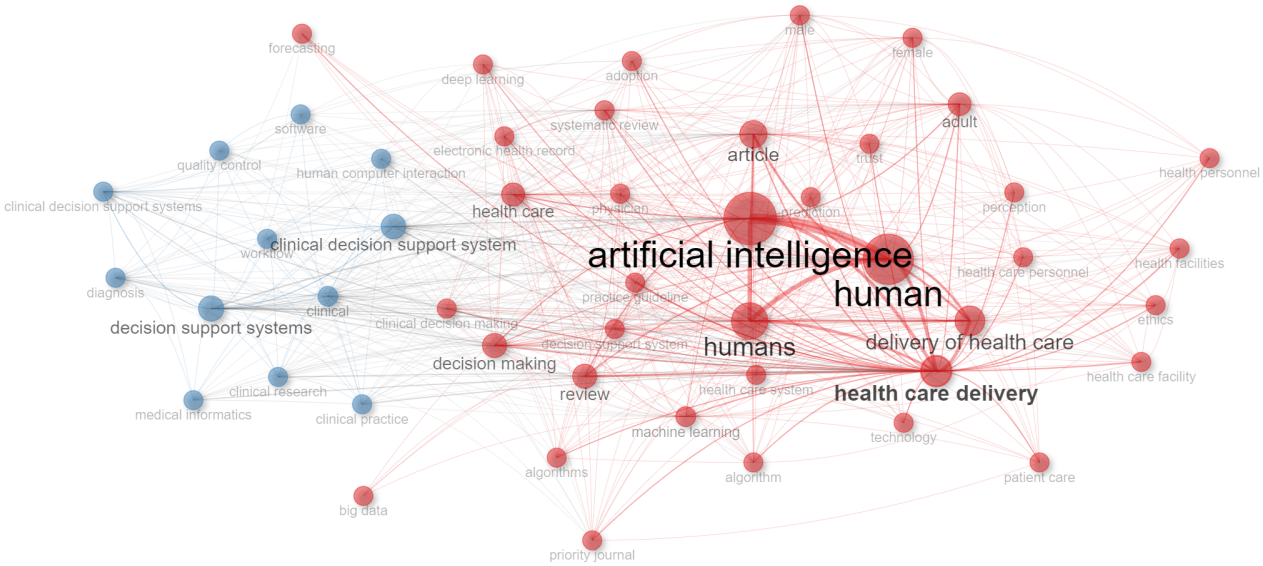
Table 3 - Source Impact

Element	h_index	TC	N.Papers	Pub_Year
Bmc Medical Informatics and Decision Making	5	608	5	2020
Npj Digital Medicine	5	249	7	2020
Digital Health	4	32	4	2021
International Journal of Environmental Research and Public Health	4	143	5	2020
Journal of Medical Ethics	4	95	4	2021
Journal of Biomedical Informatics	3	204	4	2018
Yearbook of Medical Informatics	3	247	3	2016
Bioethics	2	16	2	2021
Bmc Health Services Research	2	137	2	2018
Bmj Health and Care Informatics	2	26	2	2021

Source: Authors own work

Co-word analysis (**RQ1b**) is a bibliometric method that studies the conceptual structure of a research field by analyzing important keywords present in documents. This method creates semantic maps that help in understanding the cognitive structure of the field, as shown in *Figure 5* (Callon *et al*. 1983).

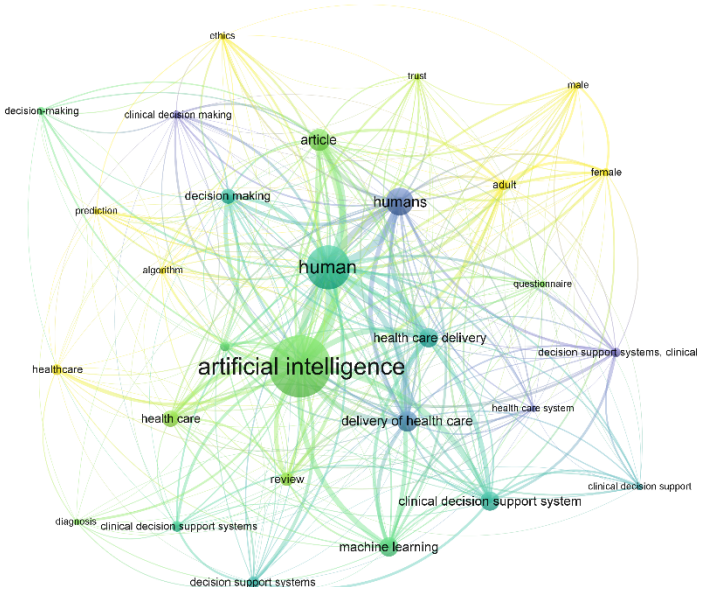
Figure 6 - Co-occurrence Network



Source: Authors own work

VOSViewer was used to create a visual summary of the main topics in the selected literature, revealing the key items and missing topics. *Figure 5* shows the predominance of discussions on the socio-technical aspects of human-machine interaction. Notably absent are concepts related to the impact of AI on healthcare professionals and patients, such as regulation, liability, cybersecurity, standards, and risk. This underscores a key assumption: AI is often treated as a technical tool for clinicians and lacks sufficient exploration from ethical, legal, and human perspectives. Less influential concepts include *ethics*, *trust*, and the *gender/age* characteristics of the actors involved, emphasizing a gap in the literature coverage (Braun *et al*., 2021; Currie and Hawk, 2021; Saheb *et al*., 2021; Kulikowski, 2022; Parakash *et al*., 2022).

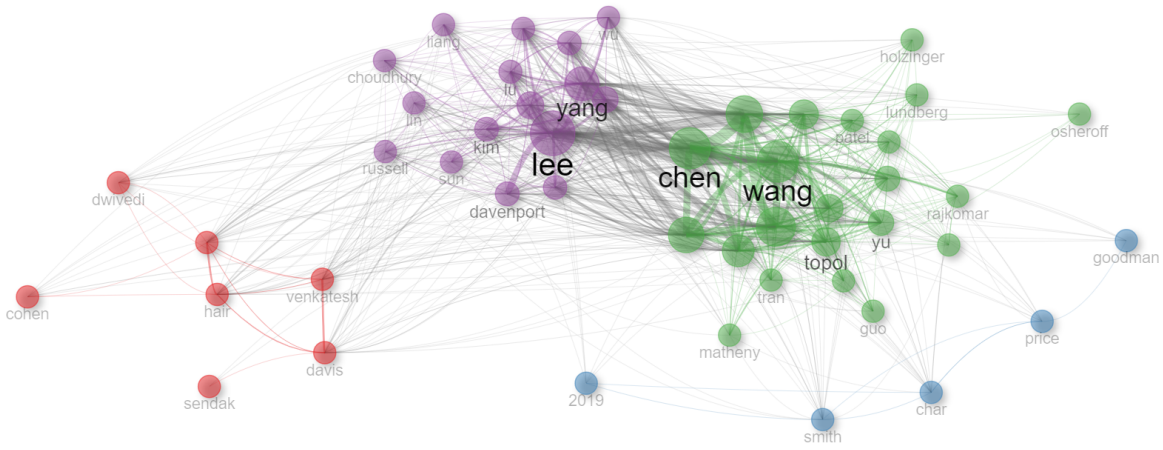
Figure 7 - Items Map



Source: Authors own work

Another bibliometric parameter is *co-author analysis network* (Glänzel, 2001; Peters and Van Raan,1991), which examines the social structure and collaboration networks of authors and their affiliations with a huge number of connections among them. This co-author analysis responds to **RQ1c**, which aimed to evaluate the way in which authors in different research areas collaborate in academic and scientific production to examine AI in healthcare as a multi-disciplinary field. As expected, the collaboration that emerges from this analysis addresses a combination of research areas, including economics, computer sciences, medicine, and neuroscience. In particular, the central cluster (in *purple*) is made up of medical and clinical experts; the *red* cluster is a more heterogeneous cluster that includes scholars from a number of areas of research, generally related to the social sciences, such as marketing and innovation experts and legal experts; the *green* cluster refers to neuroscience and technology; and the *blue* cluster includes biomedical ethics specialists.

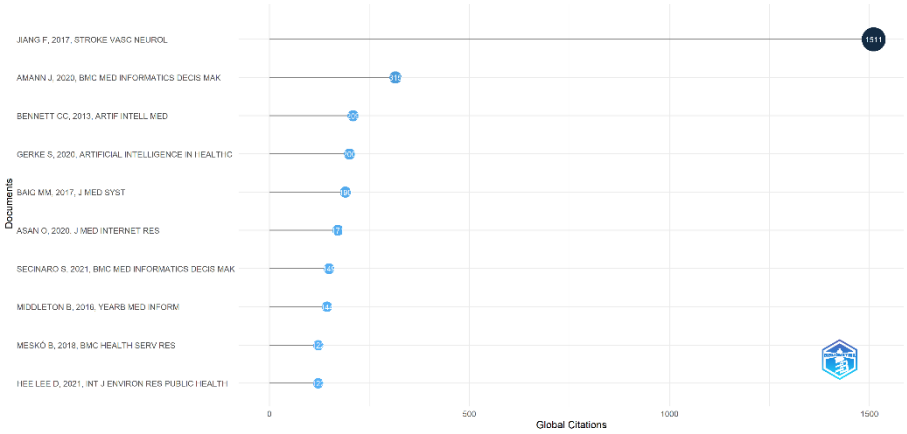
Figure 8 - Co-author analysis



Source: Authors own work

Citation analysis (Zhao and Strotmann, 2008; White and McCain, 1998; White and Griffith, 1981) is the most common form of bibliometric analysis. It counts citations to determine the similarities between documents, authors, and journals. This can allow us to identify the most important authors in a research field (**RQ1a**), and in this case has identified "*Artificial intelligence in healthcare: past, present and future*" by Jiang *et al*. (2017), as the most commonly cited paper.

Figure 9 - Most Global Cited Documents



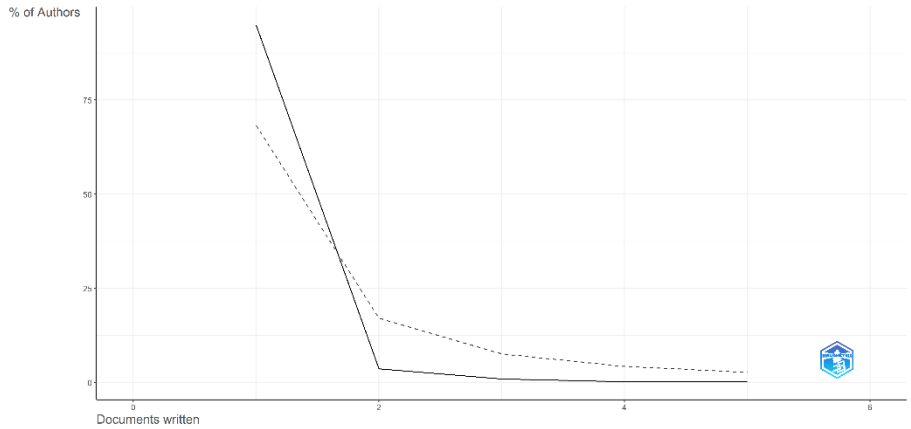
Source: Authors own work

A useful tool for tracing authorial productivity over time is Lotka's law, a mathematical formulation first introduced in 1926 to explain the frequency of publication by authors in a particular research area (Secinaro *et al*., 2021). The law states that the number of authors contributing to the research in a given period is a fraction of the number who make up a single contribution. The mathematical relationship is expressed in reverse by the formula $x^n * y_x = C$, where y_x is the number

of authors producing *x* articles in each research field. Therefore, *C* and *n* are constants that can be estimated from the calculation.

The findings of this study are consistent with those of Lotka, indicating that the average number of publications per author in a given research field is two. Furthermore, the figure illustrates the proportion of authors. Our findings (Figure 9) indicate that the research field in question is still young and growing. Approximately 75% of the authors published only their first research article. Only approximately 20% of authors have published two scientific papers.

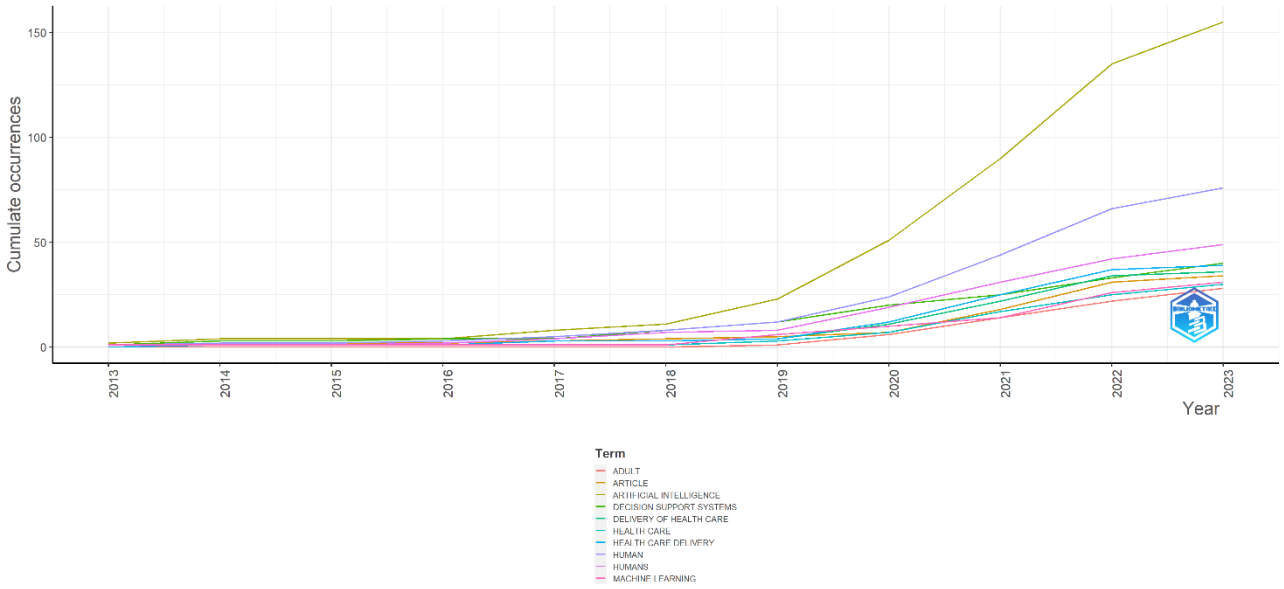
Figure 10 – Lotka's Law



Source: Authors own work

To conclude our bibliometric analysis and trace keywords occurrence over time, we provide a graph indicating the evolution of keyword use during the period under investigation (Figure 10).

Figure 11 - Most Relevant Words' Frequency over Time



Source: Authors own work

Main findings from the sample

Bibliometric analysis and SLR revealed several key themes and gaps in the literature on the implementation of AI in healthcare. One of the main themes identified is the lack of comprehensive consideration of AI systems as integrated *bundles* of algorithms, IoT, communication systems, and additional services (**RQ2**). This notable absence of a comprehensive approach to investigating the impact of AI in healthcare by recognizing the importance of these interconnected components contradicts Hypothesis 2, as shown in Table 4 (**Hp2**). The hypothesis remains unverified since the articles examined these facilities separately, with a major focus on platform and data (algorithms), and without any cohesive integration of the four facilities (Rundo *et al*., 2020; Bartels *et al*., 2022; de Hond *et al*., 2022).

The existing literature on the subject predominantly views AI as a technical tool, with limited discussion of its broader methodological impact on healthcare practices. This leads to the acceptance of **Hp1** given that only a limited number of articles address the methodological impacts of AI on healthcare procedures.

A preliminary investigation of the potential risks and adverse outcomes associated with AI adoption is evident in the literature review, particularly from the perspective of healthcare professionals. The hypothesis that academics have identified the benefits and performance improvements of adopting AI in the healthcare sector, while neglecting to address the potential concerns and harms of these systems, is accepted (**Hp3a**). This is because a review of the literature shows that only a limited number of authors have considered the potential concerns and harms

associated with AI use in comparison to the numerous studies that have highlighted its advantages. These include Antes *et al*. (2021) on the concerns and benefits of AI in healthcare, Naik *et al.* (2022) on privacy concerns and cybersecurity, Gupta *et al*. (2021) on responsible AI frameworks, Richardson *et al*. (2021) on patient concerns, Murphy *et al*. (2021) on ethical issues based on bias and accountability, and Chui *et al*. (2022) on trust issues limiting AI adoption.

The literature tends to focus more on patient-centered risk assessments (**Hp3b**), rather than considering the attitude towards AI adoption from the potential adopters' viewpoint (Esmaeilzadeh, 2020; Sunarti *et al*., 2021; Young *et al*., 2021), with relatively less focus on clinicians, who are crucial stakeholders in AI implementation (Lambert *et al*., 2023).

The literature emphasizes the importance of understanding stakeholders' perceptions and navigating the complexities of AI in healthcare. This is particularly evident in studies that examine AI adoption at the individual level, which contrasts with the recommendations set out in **Hp4**. This highlights the need for further investigation of the implications of AI in organizational contexts. It is essential to consider several factors, including the size of the organization, nature of the workflows, level of training, and security measures. This hypothesis is rejected because only a small number of articles address organizational change (Table 4). Studies addressing these issues include Asan *et al*. (2020) on organizational policies and environmental factors, Chomutare *et al*. (2022) on the potential of AI to relieve healthcare staff workload, Barda *et al*. (2020) on the implications of clinical workflows and model interpretability, and Ciecierski-Holmes *et al*. (2022) on the challenges of integrating AI in healthcare systems (Moazemi *et al*., 2023).

Table 4 – Hypotheses Evaluation Based on Systematic Review of Literature

Hypothesis	Findings (out of a total of 113 articles)	Authors	Topics
Hp1. The literature predominantly views AI as a *technical tool*, with limited discussion of its broader methodological *impacts on healthcare practices*	Among the analysed papers, the focus is primarily on specific technical applications and improvements. This is evidenced by the references included in the table, as none of them addresses broader *methodological implications*	Kerasidou (2020)	Prioritizing Human Values in AI in Healthcare
		Miller *et al*. (2020)	Digital Symptom Checkers Improve Healthcare Access
		Bates *et al*. (2021)	Data-Driven Analytics with AI for Patient Safety Improvement
		Houfani *et al*. (2021)	Precision Medicine and Improved Health Outcomes with AI
		Lee and Yoon (2021)	AI Impact on Care Delivery and Patient Lifestyle

Hp2. Authors on this topic focus on AI systems, jointly considering all structures as a bundle	In the analysed sample, only the authors cited here have considered AI systems as integrated bundles of algorithms, IoT, communication systems, and services	Rundo *et al*. (2020)	Graphical User Interfaces (GUIs) for Clinical Decision-Making
		Bartels *et al*. (2022)	Training advanced algorithms
		de Hond *et al*. (2022)	AI-Based Prediction Models
Hp3a. Academics highlight the benefits and performance improvements of adopting AI in the healthcare sector, leaving aside the potential concerns and harms of these systems	A review of the existing literature reveals a paucity of attention to the potential concerns and harms associated with the use of AI in healthcare (see the authors cited here), despite the abundance of literature addressing the advantages of AI in this field	Antes *et al*. (2021)	Concerns and Benefits of AI in Healthcare
		Gupta *et al*. (2021)	Responsible AI Frameworks for Addressing Privacy and Ethical Concerns
		Richardson *et al*. (2021)	Patient Concerns About AI in Healthcare
		Murphy *et al*. (2021)	Ethical Issues based on bias and accountability
		Chui *et al*. (2022)	Trust Issues Limit AI Adoption in Healthcare
		Naik *et al*. (2022)	Privacy concerns, Bias of Discrimination and Cybersecurity
Hp3b. Academics focus only on the risk perception of AI adoption in healthcare from the perspective of physicians and patients	The literature confirms a focus on patient-centred risk assessments, with relatively less attention accorded to the concerns of clinicians (Lambert *et al*., 2023), who are crucial stakeholders in AI implementation	Esmaeilzadeh (2020)	Consumers' Perception of AI Value and Benefits vs. Risks in Healthcare
		Sunarti *et al*. (2021)	Patients' Privacy and Autonomy Concerns
		Young *et al*. (2021)	Public Attitudes Towards AI
		Lambert *et al* (2023)	Distrust of AI Among Healthcare Professionals
Hp4. Researchers deal with AI by focusing on an *organizational level*, starting from the assumption that structural factors such as organization size, workflows, training, and security could play a key role in the adoption of AI technologies	The paucity of studies that have explored the organisational changes and disruptions that impact the integration of AI represents a notable gap in the literature. This is evidenced by the limited number of articles that have been reported in this section	Asan & Choudhury (2021)	Organizational Policies, Culture, Task Assignments, Provider Interactions, and Environmental Factors
		Barda *et al*. (2020)	Clinical Workflows and Model Interpretability in Medical ML Adoption
		Chomutare *et al*. (2022)	AI's Role in Relieving Healthcare Staff Workload
		Ciecierski-Holmes *et al*. (2022)	Disruptions in Healthcare Systems Impacting AI Integration

Source: Authors own work

The analysis highlights a lack of cross-disciplinary engagement in AI research, which is essential for realizing the full potential of algorithm implementation in healthcare (Shelmerdine *et al*., 2021).

Although some interdisciplinary collaboration exists, it still appears insufficient to address the multiple challenges and opportunities presented by AI. The impact of AI on healthcare is multifaceted, with researchers highlighting various aspects of its potential benefits and ethical considerations.

This discussion begins with studies that reveal AI's disruptive potential and then consider some beneficial AI applications. For example, Miller *et al*., (2020) and Lee and Yoon (2021) underscore the role of AI in enhancing access to healthcare through digital symptom checkers and by transforming traditional appointment systems. Ali *et al*. (2023) delve into the broader applications of AI in general healthcare, diagnostics, therapeutics, and psychiatry, emphasizing potential advantages in efficiency, precision medicine, and caregiver support (Mudgal *et al*., 2022).

The potential of AI to alleviate the workload of healthcare workers and enhance overall efficiency is a topic of the current investigation (Gerke *et al*., 2020). The emphasis is on the collaborative integration of AI with existing human resources rather than on the replacement of human workers (Meskó *et al*., 2018; Corvello *et al*., 2022). The challenges inherent in knowledge engineering, human-machine interaction, interdisciplinary collaboration, and liability concerns are acknowledged as crucial considerations in the ongoing development and implementation of AI in healthcare (Horgan *et al*., 2020; Petersson *et al*., 2022; Ferrario *et al*., 2023).

Malak *et al*. (2019) and Mehta *et al*. (2019) recognize the transformative power of AI, specifically in managing complex healthcare data, underscoring the necessity for a diverse set of tools in healthcare information management. This perspective highlights the pivotal role that AI can play in handling the intricacies of healthcare data, potentially leading to significant advancements in information management in the healthcare sector.

Table 5 - Main Authors on AI disruptive potentials' Topics

AI disruptive potential	Main topic	Authors
Main benefits - on patients	Digital Symptom Checkers Improve Healthcare Access	Miller *et al*. (2020)
	AI Impact on Care Delivery and Patient Lifestyle	Lee and Yoon (2021)
	Implementing Data-Driven Analytics with AI for Patient Safety Improvement	Bates *et al*. (2021)
Main concerns and limitations	AI in Psychiatry: DDSSs Not Ready for the Promised AI Revolution	Bertl *et al*. (2022)
	Concerns and Benefits of AI in Healthcare	Antes *et al*. (2021)
	Pitfalls and Quality Assurance in CDSS Adoption	Mahadevaiah *et al*. (2020)
	AI Limitations: Weighing Input Data, Comorbidities, Patient Contexts	Harada *et al*. (2020)
Main challenges	Challenges in AI Adoption in Europe	Horgan *et al*. (2020)
	Challenges in Implementing AI Systems in Healthcare	Petersson *et al*. (2022)

	Sociotechnical Challenges in AI Development and Adoption	Ferrario *et al*. (2023)
Potential impact - on clinical workflow	CDSS Enhancing Diagnostic Speed and Accuracy	Alabdulkarim *et al*. (2019)
	AI Impacts Decision-Making Dynamics in Healthcare	Magrabi *et al*. (2019)
	Need for Control Mechanisms in AI-Assisted CDSSs	Lysaght *et al*. (2019)
	AI Applications in Diagnostics, Therapeutics, Population Health, and Drug Discovery	Kulkov (2023)
	Prioritizing Human Values in AI in Healthcare	Kerasidou (2020)
	Efficiency Improvement in Hospitals using AI	Mudgal *et al*. (2022)
	Global Efforts in AI for Healthcare Business Ecosystems	Ismail *et al*. (2022)
	Precision Medicine and Improved Health Outcomes with AI	Houfani *et al*. (2021)
AI applications in medical world and success factors	Nurse Practitioners' Involvement in AI-Based Tools	Raymond *et al*. (2022)
	IT Transformation in Hospitals	van de Wetering and Versendaal (2018)
	AI Integration and Data Awareness	Ulloa *et al*. (2022)
	Knowledge-Based Power Structures	Sun (2021)
	Cognitive Computing	Srivani *et al*. (2023)
	Quality-in-Use Characteristics for CDSS	Souza-Pereira *et al*. (2021)
	AI Research Growth in Healthcare	Shelmerdine *et al*. (2021)
	AI Applications in Healthcare	Senbekov *et al*. (2020)
	AI's Impact on Healthcare and Future Directions	Secinaro *et al*. (2021)
	Building CDSSs	Raja and Asghar (2020)
	Knowledge Embodiment in AI Robotic Systems	Pee *et al*. (2019)
Organizational Changes	Adoption of Wearable Patient Monitoring (WPM) by Clinicians	Baig *et al*. (2017)
	Disruptions in Healthcare Systems Impacting AI Integration	Ciecierski-Holmes *et al*. (2022)
	Alignment of CDSS with Clinical Workflows and Model Interpretability in Medical ML Adoption	Barda *et al*. (2020)
	AI's Role in Relieving Healthcare Staff Workload	Chomutare *et al*. (2022)
AI vs Clinicians: Performances	AI Outperforming Humans in Medical Processes	Ali *et al*. (2023)
	AI complements Medical Professionals	Meskó *et al*. (2018)
	AI Performance Compared to Clinicians	Shen *et al*. (2019)
	Challenges in Medical Knowledge Engineering and Human-Machine Interaction	Johnson *et al*. (2014)
	AI Matching Healthcare Professionals in Diagnostics	Biller-Andorno *et al*. (2021)
	AI potential to improve Healthcare Professionals' Activities	Gerke *et al*. (2020)
	AI Simulation for Optimal Treatment Decisions in Complex Healthcare Systems	Bennet and Hauser (2013)
	CDSSs: increase and not replacement of Human Expertise	Middleton *et al*. (2016)
Shared Knowledge	Collaborative Science Teams (CSTs)	Wilson *et al*. (2021)
	Automatic Information Fusion for Handling Big Healthcare Data in Smart Healthcare	Chen *et al*. (2023)
	Collaboration Between Research and Clinic in AI-Decision Support System Use	Braun *et al*. (2021)

	Impact of Leadership, Workflow Understanding, and Proprietary Implementations	Greenes *et al*. (2018)
Data accessibilty and utilization	Data Utilization Potential in Healthcare	Mehta *et al*. (2019)
	Hurdles in AI Adoption: Data Accessibility and Regulatory Complexities	Kasparick *et al*. (2019)
	AI Systems Training with Clinical Data	Jiang *et al*. (2017)
	Diverse Tools Needed for Complex Healthcare Data	Malak *et al*. (2019)

Source: Authors own work

The integration of AI in healthcare presents promising advancements but also introduces *ethical* and *legal* challenges that extend beyond technological concerns. Key hurdles include safety, transparency, and algorithmic fairness (Bates *et al*., 2021). Guidelines for transparency and safety are necessary to address concerns regarding AI-induced harm and promote trust (Lysaght *et al*., 2019). Moreover, the *black-box* problem, which hinders understanding and trust, calls for solutions that incorporate explainable AI (XAI) (Amann *et al*., 2020; Jin *et al*., 2021; Nazar *et al*., 2021; Wadden, 2021; Yang, 2022; Albahri *et al*., 2023; Benzinger *et al*., 2023; Bienefeld *et al*., 2023; Messeni Petruzzelli *et al*., 2023; Zhang and Zhang, 2023). The ethical debates surrounding healthcare focus on the principles of fair decision-making, validation of health support systems, and ability of professionals to interpret the recommendations provided by AI (Magrabi *et al*., 2019; Gurupur *et al*., 2020; Murphy *et al*., 2021; Kulkov, 2023).

Ethical imperatives are demanded owing to concerns regarding privacy, bias, discrimination, and cybersecurity (Naik *et al.,* 2022). Legal considerations include licensing, standardization, and funding for AI applications in healthcare (Wani *et al*., 2022). Moreover, clinicians' responsibility for patient harm caused by AI recommendations raises ethical and legal questions that require shared responsibility, oversight, and control systems (Bhardwaj, 2022). To ensure accountability for patient harm, it is essential to establish governance principles, legal guidelines, and a comprehensive ethical framework at the national and EU levels (Taeihagh, 2021; Bedoya *et al*., 2022).

The analyzed literature also emphasizes the critical importance of implementing AI in a responsible manner within the healthcare sector. This entails addressing several key issues such as *ensuring diversity in data*, *fostering collaboration* and *building trust*. This approach is crucial for cultivating favorable attitudes among healthcare professionals (Fritsch *et al*., 2022; Daich Varela *et al*., 2023). Furthermore, the literature indicates that clinicians' attitudes towards AI are significantly shaped by patient concerns, underscoring the necessity for a holistic comprehension of patient perspectives (Richardson *et al*., 2021, 2022).

Furthermore, this study examines the role of attitude, cognitive biases, and heuristics in influencing decision makers in uncertain processes, as discussed by Chua *et al*. (2021) and Blumenthal-Barby and Krieger (2015).

Table 6 - Main Authors on Ethical and Legal Concerns' Topics

Ethical and Legal Issues	Main topic	Authors
Ethical challenges, informed consent, fairness	Ethical Categories in AI Ethics Discourse	Saheb *et al*. (2021)
	Ethical Challenges in AI-Based Clinical Decision Support	Rogers *et al*. (2021)
	Ethical Considerations in AI-CDS for Children	Ramgopal *et al*. (2023)
	Privacy concerns, Bias of Discrimination and Cybersecurity	Naik (2022)
	Ethical Considerations in Nuclear Medicine	Currie and Hawk (2021)
Safety, transparency, potential algorithm failures	Unexplainable AI and Its Impact on Clinical Decision-Making	Kulikowski (2022)
	Transparency and Explainability	Valente *et al*. (2022)
	"Black-Box" Problem	Wadden (2021)
	Licensing, Standardization, Financing	Wani *et al*. (2022)
Need for a regulatory Framework and a standardized discipline	Governance Principles and Best Practices for AI/ML-Based CDSSs Development	Bedoya *et al*. (2022)
	Comprehensive Ethical Framework for AI in Healthcare	Prakash *et al*. (2022)
	Need for a technology-Based Healthcare Framework	Yoon and Lee (2019)
Shared responsibility in AI development and application	Responsible AI framework for healthcare	Siala and Wang (2022)
	ML Algorithm Responsibility and Cost-Benefit Considerations in providing High-Value Healthcare	Bhardwaj (2022)
	Responsible AI Frameworks for Addressing Privacy and Ethical Concerns	Gupta *et al*. (2021)
	Awareness of AI's Role in Treatment Decisions and Legal Considerations	Amann *et al*. (2020)
	Ethical Issues based on bias and accountability	Murphy *et al*. (2021)
Training advanced algorithms with healthcare data	Usability Testing and User-Centered Design for Medical Devices	Asan and Choudhury (2021)
	Quality Control Measures for AI/ML-CDS Models Transitioning to Clinical Care	Bartels *et al*. (2022)
	Clinical Ethics in AI Lacks Focus on Justice, Explicability, and Human-Machine Interaction	Benzinger *et al*. (2023)
	Interpretability Central to AI Integration in Healthcare	Moazemi *et al*. (2023)
	Framework and Recommendations for Bridging the Gap Between Clinicians and Developers in Explainable AI	Bienefeld *et al*. (2023)
	Explainable AI (XAI)	Yang (2022)
	Back-Propagation-Based Interpretability Methods for Deep Learning Models	Jin *et al*. (2021)
	Understanding Explainable AI (XAI) in HCI	Nazar *et al*. (2021)
	Rigorous Validation for AI-Based Decision Support Systems	Gurupur *et al*. (2020)
	Guidelines and Quality Criteria for AI-Based Prediction Models	de Hond *et al*. (2022)

Source: Authors own work

Trust is a pivotal factor in the acceptance of AI by clinicians (Del Giudice *et al*., 2023; Magni *et al*., 2023b). Factors such as confidence-building measures, human trust in computer systems, social

influence (Sun, 2021), and design considerations may influence AI adoption (Asan *et al.*, 2020; Shinners *et al*., 2020, 2022; Cheng *et al*., 2022; Rojas *et al*., 2023). As discussed by some scholars, resistance to the recommendations of CDSSs arises from concerns regarding trust, workflow, and individual perceptions. User-friendly interfaces, transparency, and optimization of decision-support alerts are crucial for the adoption of CDSSs (Greenes *et al*., 2018; Shen *et al*., 2019). Accordingly, a balanced approach to trust is necessary (Liang and Xue, 2022). Glick *et al*. (2022) and Naiseh *et al.* (2023) explore collaborative AI decision-making while preserving clinicians' autonomy. Negative public perception poses a barrier (Castagno and Khalifa, 2020), necessitating alignment with clinician needs, patient views, and collaboration with healthcare professionals and technology developers for successful AI implementation in healthcare workflows (Yoon and Lee, 2019; Wolff *et al*., 2021).

Furthermore, digitalization of the healthcare supply chain has become an urgent necessity, particularly in the context of global uncertainties and pandemics. Bag *et al*. (2023) investigate the antecedents of big data analytics (BDA) (Chen *et al*., 2023) and AI technology-based collaborative platforms for enhancing absorptive capacity in omni-channel healthcare processes. These findings indicate that such platforms can significantly enhance organizational performance by enabling healthcare supply chains to assimilate, transfer, and exploit critical information from large datasets, thereby delivering innovative performance. The integration of BDA-AI technologies in the healthcare supply chain serves to illustrate the transformative potential of digital tools in optimizing healthcare delivery.

Considering the significant impact of these systems on the division of labour and roles, some authors have directed their attention towards organizational factors, including policies, culture, and environmental elements, which may also influence the acceptance and adoption of AI in the healthcare sector (Asan *et al*., 2020).

Table 7 – Main Authors on Attitude and Impact on Workflows' Topics

Attitude & Impact on Workflows	Main topic	Authors
Factors influencing AI adoption	Factors Impacting Trust in AI Adoption Among Healthcare Professionals	Asan and Choudhury (2020)
	Factors Influencing AI Adoption in Healthcare	Khanijahani *et al*. (2022)
	Factors Affecting Clinicians' Intent to Use AI-Based Systems	Choudhury (2022)
	Healthcare Professionals' Trust in AI	Shinners *et al*. (2020, 2022)
	Trust-Building Measures	Wolff *et al*. (2021)
	Clinicians' Expectancy, Trust, and Perceptions in AI Adoption	Choudhury *et al*. (2022)

	Influence of Human-Computer Trust and Social Influence in AI Adoption in Healthcare	Cheng *et al*. (2022)
	Trust Issues Limit AI Adoption in Healthcare	Chui (2022)
	Trust Calibration in Collaborative AI Decision-Making	Naiseh (2023)
	Pessimistic Public Perception of Medical AI and Its Impact on Integration	Castagno and Khalifa (2020)
	Trust-Building Attributes in Human-AI Interactions	Rojas *et al*. (2023)
	Physician Resistance to CDSS Recommendations	Liang and Xue (2022)
	Trust in AI Systems: Factors Including Organizational Policies, Culture, Task Assignments, Provider Interactions, and Environmental Factors	Asan *et al*. (2020)
Fully automatized systems	Graphical User Interfaces (GUIs) for Clinical Decision-Making	Rundo *et al*. (2020)
Risk Perception	Perception of AI Value and Benefits vs. Risks in Healthcare	Esmaeilzadeh (2020)
	Distrust of AI Among Healthcare Professionals	Lambert *et al*. (2023)
	Privacy and Autonomy Concerns	Sunarti (2021)
	Public Attitudes Towards AI	Young *et al*. (2021)
Impact on clinical workflows	General practitioners benefit from CDSSs	Westerbeek *et al*. (2021)
	Challenges in Implementing CDSSs	O' Sullivan *et al*. (2014)
	AI Integration in Health Systems	Kashyap *et al*. (2021)
	AI Input Affecting Clinical Decision-Making	Glick *et al*. (2022)
Relationship between AI providers and healthcare clinicians	Responsible AI Implementation: Addressing Data Privacy, Bias, Expertise, Collaboration, and Trust-Building	Alowais *et al*. (2023)
	AI Adoption Challenges: Data Diversity, Expert Collaboration, User-Friendly Interfaces	Daich Varela *et al*. (2023)
Patients' Perceptions	Patient Concerns About AI in Healthcare	Richardson *et al*. (2021, 2022)
	Patient Views on AI in Healthcare	Fritsch *et al*. (2022)

Source: Authors own work

These findings underscore the transformative potential of AI in healthcare, highlighting its ability to enhance clinical decision-making, improve patient outcomes, and increase efficiency in healthcare delivery through Clinical Decision Support Systems (CDSSs). However, when examining the developing field of AI in healthcare, it is crucial to acknowledge the complex nature of AI technologies and their significant impact on *human–technology interaction* (Caputo *et al*., 2023b; Chin *et al*., 2024). The reviewed academic literature suggests that while AI offers promising benefits, its successful integration into healthcare requires careful consideration of its social and technical aspects to ensure that technological advancements align with human values and ethical standards (Kerasidou, 2020; Kashyap *et al*., 2021, Souza-Pereira *et al*., 2021). The synthesis of bibliometric insights highlights a burgeoning field of AI in healthcare, characterized by rapid growth and multidisciplinary collaboration; however, there is a significant gap in the literature concerning the

humanistic aspects of AI integration, such as ethical considerations, regulatory challenges (Kasparick *et al*., 2019), and the nuances of human-machine interaction and risk management. The current focus on AI as a tool in the clinician's toolkit may also overlook its transformative potential and the significant ethical, legal, and interpersonal implications associated with it (Del Giudice *et al*., 2023; Magni *et al*., 2023b).

Despite the advancements and potential benefits described by Miller *et al.* (2020) and Lee and Yoon (2021), which include enhanced *access to healthcare* and *efficiency improvements* through AI-driven diagnostics and patient care (Alabdulkarim *et al*., 2019), our analysis has uncovered a critical oversight: there appears to be a lack of comprehensive engagement with how these technologies affect the roles, responsibilities, and experiences of healthcare professionals and patients (Biller-Adorno *et al*., 2021; Del Giudice *et al*., 2023; Magni *et al*., 2023a).

In particular, our co-author analysis demonstrates a lively multidisciplinary collaboration that covers economics, computer science, medicine, and neuroscience. This interdisciplinary approach highlights the recognition of AI in healthcare as a domain that goes beyond traditional boundaries, requiring a harmonious fusion of technical expertise and clinical insight. While collaboration metrics suggest a healthy cross-pollination of ideas, they also indicate a missed opportunity to engage more deeply with the humanistic and ethical dimensions of technology deployment in healthcare environments (Asan *et al.*, 2020; Shinners *et al*., 2020; 2022). Even the citation analysis indicates a growing interest in the practical applications of AI in healthcare. However, research addressing the ethical and human-centric considerations of such technologies is lacking. This oversight contradicts the general discourse on the need for a more nuanced understanding of AI's role in healthcare, which fully acknowledges the complexities of human–machine symbiosis (Caputo *et al*., 2019b).

As our review reveals, the intersection of AI with healthcare elicits a spectrum of reactions from the medical community, ranging from optimism about its potential to augment human capabilities to apprehension regarding the ethical dilemmas it poses. Drawing on insights from Del Giudice *et al*. (2023a) and Magni *et al*. (2023a), it is imperative that future research endeavors delve deeper into the implications of AI for the fabric of healthcare delivery, exploring not only the technical efficiencies it promises but also the ethical, legal, and human concerns it engenders.

In light of this, our analysis strongly advocates for a shift in scholarly focus towards a more holistic examination of AI in healthcare. Such an approach is crucial for fostering a healthcare ecosystem, where technology serves not only as a facilitator of clinical tasks but also as a catalyst for enhancing the human aspects of care delivery, aligning with the principles of Society 5.0, and advocating a heartily collaborative approach that prioritizes the well-being and dignity of all stakeholders involved (Caputo *et al*., 2019a; Ismail *et al*., 2022; Del Giudice *et al*., 2023b; Magni *et al*., 2023b).

Discussion and Conclusions

We have undertaken a comprehensive study utilizing bibliometric analysis to identify scholars, collaborations, and main topics in the research area of AI in healthcare. Our findings indicate that this field is rapidly expanding and attracting an increasing number of researchers. The analysis employed both a quantitative approach to examine bibliometric variables and a qualitative approach to study recurring keywords. Three main findings were obtained:

- The study found a *regulatory gap* in the medical application of AI, suggesting the need for a legal and ethical normative framework that might address issues of responsibility in the event of patient harm.
- Healthcare professionals show *limited willingness to adopt* AI systems, emphasizing the importance of legal and ethical frameworks in alleviating concerns about individual responsibility in the event of patient harm.
- Despite these challenges, there is unprecedented *potential for healthcare transformation* through Clinical Decision Support Systems (CDSSs) (O' Sullivan *et al*., 2014).

From the perspective of Prospect Theory, the findings suggest that healthcare professionals may not readily adopt CDSSs without a clear legal and ethical framework that can mitigate perceived risks, especially in terms of individual responsibility for harm caused by algorithmic recommendations.

The literature review primarily emphasized the positive impact of AI in healthcare, focusing on gains and innovations. However, the loss frame, which tackles potential drawbacks and risks, remains only marginally addressed. We highlight the need for a systematic examination of users' perceptions of risk, which underscores the importance of regulatory guidelines. Since regulation is crucial for governing the use, impact, and suitability of AI facilities in healthcare, we call for national and supranational regulatory measures, particularly to deal with issues of risk regulation in AI-involved business processes within healthcare systems.

As previously stated, this study emphasizes the need for a deeper exploration of concepts such as *risk perception*, *cybersecurity*, and the role of AI-embedded systems as decision makers. In particular, those with expertise in the field advocate a comprehensive approach to understanding the psychological impact of AI on healthcare practitioners.

Theoretically, this research contributes to the advancement of the understanding of the role of AI in healthcare by applying Prospect Theory to elucidate the psychological factors influencing the adoption of AI in clinical decision-making. The analysis also identifies a significant gap in the existing literature regarding the comprehensive integration of AI systems and the broader methodological impacts on healthcare practices. It contributes to the discourse on risk perception and the ethical implications of AI, advocating a nuanced understanding of stakeholders' concerns.

The originality of this research lies in its interdisciplinary approach, which integrates perspectives from *algorithms*, *IoT*, *communication systems*, and *services* to evaluate AI systems as integrated bundles.

Practically, this study underscores the importance of considering organizational factors in the successful adoption of AI technologies. It offers valuable insights for healthcare practitioners, policymakers, and technology developers by emphasizing the need for a holistic approach to AI integration that aligns with the ethical, legal, and humanistic dimensions of healthcare delivery.

Limitations and Future Research

Despite its contributions, this study has certain limitations. The reliance on bibliometric methods may introduce biases related to source selection and analysis. Additionally, the focus on specific databases (Scopus, Web of Science, and PubMed) may have excluded relevant literature from other sources. The scope of this study is limited to the examination of AI in healthcare without a deep dive into specific subfields or technologies, which could be explored in future research.

It would be beneficial for future research to investigate the long-term impact of AI on healthcare outcomes, considering both immediate and long-term effects. A more nuanced understanding of the specific barriers and facilitators of AI adoption in diverse healthcare settings can be obtained through further investigation. Moreover, interdisciplinary studies involving stakeholders from a range of disciplines, including ethics, law, and social sciences, are essential for the development of a comprehensive framework for the responsible implementation of AI in healthcare. A more thorough exploration of the patient perspective and an examination of the socioeconomic implications of AI adoption in healthcare can further enhance the understanding and application of AI technologies. Further research is required in this area. To this end, in the next phase of our research, we plan to use the Technology Acceptance Model (TAM) and the Corporate Innovation Matrix (CIM) to analyze the impact of AI on micro- and macro-business processes. Additionally, we intend to employ Prospect Theory (PT) and Motivational Model (MM) to investigate the psychological effects on practitioners and potential users, thereby enhancing understanding and guiding responsible AI development.

Moreover, the implementation of effective Knowledge Management (KM) practices will prove instrumental in addressing these limitations. The deliberate concealment of information, or "knowledge hiding", may manifest as a response to AI-driven job security concerns, particularly in high-stakes healthcare settings. This behaviour may be intensified by the technological turbulence generated by AI, resulting in resistance to the sharing of knowledge that could otherwise facilitate AI processes. Conversely, the promotion of a culture of knowledge sharing and implementation of robust KM systems may prove to be an effective means of addressing these issues (Nguyen *et al*., 2024).

The creation of a secure and collaborative environment can allay concerns about job displacement, thereby facilitating the integration of AI in a more seamless manner. It would be beneficial for future research to investigate how KM practices can be employed to effectively counteract the phenomenon of knowledge hiding, thereby optimizing the potential of AI in enhancing healthcare outcomes while addressing the concerns of clinicians regarding role security.

Theoretical and practical implications

To bridge the gap between theoretical insights and practical applications, this study underscores the transformative potential of AI in healthcare, highlighting the ethical, legal, and humanistic considerations that will be needed if it is to be integrated.

From a theoretical standpoint, while acknowledging the utility of the framework of Prospect Theory (PT) in examining psychological risk factors across diverse contexts, it would be beneficial for scholars in this field to enhance the framework's utility and applicability by developing both qualitative and quantitative research methods to measure individual risk perception (sensitivity to risk) and risk-taking (propensity to engage in risky activities). To date, this theory, which is typically regarded as the theoretical foundation of behavioural economics and decision making, has been employed to propose an alternative to the Expected Utility Theory (EUT) as a descriptive model of decision making under conditions of risk and uncertainty. The assumptions of PT consider the gain and loss frames in monetary terms, and do not consider individuals' risk perception in a domain-specific context.

Given these limitations, we would propose a hybrid approach that integrates the PT framework with the Domain-Specific Risk-Taking (DoSpeRT) methodology (Blais and Weber, 2006). This integrating step would assure researchers to consider the risk attitudes of individuals across different categories of risks (such as ethical, financial, health/safety, recreational, and social) in a domain-specific context.

Indeed, starting from the same assumptions as PT – according to which risk attitude can be defined as the '*chosen response to uncertainty that matters, driven by perception*' (Hillson and Murray-Webster, 2007) – the DoSpeRT methodology has the potential to allow scholars to resolve the two main problems that emerge when calculating risk attitudes using the EUT framework: 1) it has been demonstrated that risk attitudes are not consistent across methodologies; 2) individuals do not exhibit consistent risk-taking behaviour in different situations, even when the same methodology is used (Loomes and Pogrebna, 2014).

PT does not directly address the domain differences. However, there are two potential causes of the instability of risk preferences. The first is the concept of *framing*, which can alter personal

benchmarks and subsequently influence one's approach to risk (the framing effect). Second, loss aversion is experienced in different domains. Instead, the DoSpeRT methodology has the potential to enhance our understanding by accounting for how *framing* and *loss aversion* influence risk attitudes differently across various domains.

In summary, it is recommended that a domain-specific methodology be employed to assess and depict the risk propensity associated with AI adoption in healthcare due to the context-specific issues that need to be considered, such as the irreversibility of a medical decision and the precarity of patients' health status. Risk propensity can be measured in terms of Willingness to Accept (WTA) an AI solution and Willingness to Pay (WTP) for it. For instance, healthcare professionals or patients could be queried regarding their WTA regarding the application of AI systems in accordance with their risk perception and risk-taking attitude. Conversely, healthcare structures and patients could be asked to indicate their WTP for the diversification of services provided and for receiving a medical performance with "zero human errors".

In conclusion, this study contributes to the growing body of literature on the adoption and integration of AI in healthcare with a particular focus on CDSSs. By applying Prospect Theory to the decision-making processes of healthcare professionals, this research contributes to the advancement of understanding of cognitive biases, such as risk aversion and loss aversion, in the context of AI adoption. The findings demonstrate that whereas previous studies have predominantly focused on the technical and operational aspects of AI, this work primarily elucidates the psychological and behavioural dimensions that influence AI adoption. This study contributes to the broader field of behavioural economics in healthcare innovation by illustrating how cognitive biases influence decision making in high-risk medical contexts. Moreover, the work introduces the concept of cognitive debiasing through AI, offering a novel theoretical framework for understanding how AI can counteract the biases that may hinder its adoption. This theoretical expansion not only addresses a gap in the existing literature but also provides a foundation for future research exploring the interplay between AI technologies and human decision-making in complex healthcare settings.

From a practical standpoint, integrating AI into healthcare enhances knowledge management (KM) frameworks by streamlining the acquisition and application of specific capabilities and skills within the sector (Botega and da Silva, 2020; Issac *et al*., 2022; Del Giudice *et al*., 2023a). Indeed, integrated AI platforms assist in the efficient handling and sharing of information, which are crucial for making informed decisions and improving healthcare outcomes (Iaia *et al*., 2023; Toscani, 2023; Serenko, 2024). This technology not only automates data processing and pattern learning but also supports continuous professional development by providing practitioners with actionable insights.

The synergy between AI and KM boosts the efficiency, adaptability, and intelligence of the healthcare system, fundamentally transforming care delivery and medical research (Dal Mas *et al*., 2020).

However, the widespread adoption of AI technologies would challenge healthcare professionals to improve their digital competencies. Such a shift would necessitate updates in medical education and professional development to equip healthcare workers with skills in data analysis, machine learning, and digital ethics, which are essential for leveraging AI in diagnostics, treatment planning, and patient monitoring (De Waal *et al*., 2016; Baig *et al*., 2017; Papa *et al*., 2020; Mele *et al*., 2023). Consequently, there is a clear need to adapt *career development programs* and *educational curricula* to prepare healthcare professionals for a future in which AI plays a critical role in healthcare systems (Messeni Petruzzelli *et al*., 2010). This adaptation would aim not only to improve patient care but also to ensure that healthcare professionals remain at the forefront of medical innovation and are able to effectively manage the complexity of modern healthcare environments.

References

Al Hawamdeh, N. (2023). Does humble leadership mitigate employees' knowledge-hiding behaviour? The mediating role of employees' self-efficacy and trust in their leader. *Journal of Knowledge Management*, 27(6), 1702-1719. https://doi.org/10.1108/JKM-05-2022-0353

Alabdulkarim, A., Al-Rodhaan, M., Tian, Y., & Al-Dhelaan, A. (2019). A privacy-preserving algorithm for clinical decision-support systems using random forest. *Computers, Materials and Continua*, *58*(2), 585–601. https://doi.org/10.32604/cmc.2019.05637

Albahri, A. S., Duhaim, A. M., Fadhel, M. A., Alnoor, A., Baqer, N. S., Alzubaidi, L., ... & Deveci, M. (2023). A systematic review of trustworthy and explainable artificial intelligence in healthcare: Assessment of quality, bias risk, and data fusion. *Information Fusion*, 96, 156-191. https://doi.org/10.1016/j.inffus.2023.03.008

Ali, O., Abdelbaki, W., Shrestha, A., Elbasi, E., Alryalat, M. A. A., & Dwivedi, Y. K. (2023). A systematic literature review of Artificial Intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities. *Journal of Innovation and Knowledge*, *8*(1). https://doi.org/10.1016/j.jik.2023.100333

Alowais, S. A., Alghamdi, S. S., Alsuhebany, N., Alqahtani, T., Alshaya, A. I., Almohareb, S. N., Aldairem, A., Alrashed, M., Bin Saleh, K., Badreldin, H. A., Al Yami, M. S., Al Harbi, S., & Albekairy, A. M. (2023). Revolutionizing healthcare: The role of Artificial Intelligence in clinical practice. *BMC Medical Education*, *23*(1). https://doi.org/10.1186/s12909-023-04698-z

Amann, J., Blasimme, A., Vayena, E., et al. (2020). Explainability for artificial intelligence in healthcare: A multidisciplinary perspective. *BMC Medical Informatics and Decision Making*, 20, 310. https://doi.org/10.1186/s12911-020-01332-6

Ammirato S., Felicetti A.M., Rogano D., Linzalone R., Corvello V. (2022). Digitalizing the Systematic Literature Review process: the MySLR platform, *Knowledge Management Research & Practice*, DOI: 10.1080/14778238.2022.2041375

Antes, A. L., Burrous, S., Sisk, B. A., Schuelke, M. J., Keune, J. D., & DuBois, J. M. (2021). Exploring perceptions of healthcare technologies enabled by Artificial Intelligence: An online, scenario-based survey. *BMC Medical Informatics and Decision Making*, *21*(1). https://doi.org/10.1186/s12911-021-01586-8

Arain, G. A., Hameed, I., Khan, A. K., Strologo, A. D., & Dhir, A. (2022). How and when do employees hide knowledge from co-workers?. *Journal of Knowledge Management*, 26(7), 1789-1806. https://doi.org/10.1108/JKM-03-2021-0185

Arias-Pérez, J., & Vélez-Jaramillo, J. (2022). Understanding knowledge hiding under technological turbulence caused by artificial intelligence and robotics. Journal of Knowledge Management, 26(6), 1476-1491. https://doi.org/10.1108/JKM-01-2021-0058

Asan, O., & Choudhury, A. (2021). Research trends in Artificial Intelligence applications in human factors health care: Mapping review. *JMIR Human Factors*, *8*(2). https://doi.org/10.2196/28236

Asan, O., Bayrak, A. E., & Choudhury, A. (2020). Artificial Intelligence and Human Trust in Healthcare: Focus on Clinicians. *Journal of Medical Internet Research*, *22*(6). https://doi.org/10.2196/15154

Babushkina, D., & Votsis, A. (2022). Epistemo-ethical constraints on AI-human decision making for diagnostic purposes. *Ethics and Information Technology*, 24(2), 22. https://doi.org/10.1007/s10676-022-09629-y

Bag, S., Dhamija, P., Singh, R. K., Rahman, M. S., & Sreedharan, V. R. (2023). Big data analytics and artificial intelligence technologies based collaborative platform empowering absorptive capacity in health care supply chain: An empirical study. Journal of Business Research, 154, 113315. https://doi.org/10.1016/j.jbusres.2022.113315

Baig, M. M., GholamHosseini, H., Moqeem, A. A., Mirza, F., & Lindén, M. (2017). A Systematic Review of Wearable Patient Monitoring Systems – current challenges and opportunities for clinical adoption. *Journal of Medical Systems*, *41*(7). https://doi.org/10.1007/s10916-017-0760-1

Bakkalbasi, N., Bauer, K., Glover, J., and Wang, L. (2006) Three options for citation tracking: Google Scholar, Scopus, and Web of Science. *Biomed. Digit. Libr*. 3, 7. https://doi.org/10.1186/1742-5581-3-7

Barda, A. J., Horvat, C. M., & Hochheiser, H. (2020). A qualitative research framework for the design of user-centered displays of explanations for Machine Learning model predictions in healthcare. *BMC Medical Informatics and Decision Making*, *20*(1). https://doi.org/10.1186/s12911-020-01276-x

Bartels, R., Dudink, J., Haitjema, S., Oberski, D., & van 't Veen, A. (2022). A Perspective on a Quality Management System for AI/ML-Based Clinical Decision Support in Hospital Care. *Frontiers in Digital Health*, *4*. https://doi.org/10.3389/fdgth.2022.942588

Bates, D. W., Levine, D., Syrowatka, A., Kuznetsova, M., Craig, K. J. T., Rui, A., Jackson, G. P., & Rhee, K. (2021). The potential of Artificial Intelligence to improve patient safety: A scoping review. *Npj Digital Medicine*, *4*(1). https://doi.org/10.1038/s41746-021-00423-6

Bedoya, A. D., Economou-Zavlanos, N. J., Goldstein, B. A., Young, A., Jelovsek, J. E., O'brien, C., Parrish, A. B., Elengold, S., Lytle, K., Balu, S., Huang, E., Poon, E. G., & Pencina, M. J. (2022). A framework for the oversight and local deployment of safe and high-quality prediction models.

Journal of the American Medical Informatics Association, *29*(9), 1631–1636. https://doi.org/10.1093/jamia/ocac078

Bennett, C. C., & Hauser, K. (2013). Artificial intelligence framework for simulating clinical decision-making: a Markov decision process approach. *Artificial intelligence in medicine*, *57*(1), 9–19. https://doi.org/10.1016/j.artmed.2012.12.003

Benzinger, L., Ursin, F., Balke, W. T., et al. (2023). Should artificial intelligence be used to support clinical ethical decision-making? A systematic review of reasons. *BMC Medical Ethics, 24*, 48. https://doi.org/10.1186/s12910-023-00929-6

Bertl, M., Ross, P., & Draheim, D. (2022). A survey on AI and decision support systems in psychiatry – Uncovering a dilemma. *Expert Systems with Applications*, *202*. https://doi.org/10.1016/j.eswa.2022.117464

Bhardwaj, A. (2022). Promise and Provisos of Artificial Intelligence and Machine Learning in Healthcare. *Journal of Healthcare Leadership*, *14*, 113–118. https://doi.org/10.2147/JHL.S369498

Bhatti, S. H., Hussain, M., Santoro, G., & Culasso, F. (2023). The impact of organizational ostracism on knowledge hiding: analysing the sequential mediating role of efficacy needs and psychological distress. *Journal of Knowledge Management*, 27(2), 485-505. https://doi.org/10.1108/jkm-03-2021-0223

Bienefeld, N., Boss, J. M., Lüthy, R., Brodbeck, D., Azzati, J., Blaser, M., Willms, J., & Keller, E. (2023). Solving the explainable AI conundrum by bridging clinicians' needs and developers' goals. *NPJ digital medicine*, *6*(1), 94. https://doi.org/10.1038/s41746-023-00837-4

Biller-Andorno, N., Ferrario, A., Joebges, S., Krones, T., Massini, F., Barth, P., Arampatzis, G., & Krauthammer, M. (2021). AI support for ethical decision-making around resuscitation: Proceed with care. *Journal of Medical Ethics*, *48*(3), 175–183. https://doi.org/10.1136/medethics-2020-106786

Blais, A.-R., & Weber, E. U. (2006). A Domain-Specific Risk-Taking (DOSPERT) scale for adult populations. *Judgement and Decision Making*, 1(1), 33–47. https://doi.org/10.1017/S1930297500000334

Blumenthal-Barby, J. S., & Krieger, H. (2015). Cognitive biases and heuristics in medical decision making: A critical review using a systematic search strategy. *Medical Decision Making*, *35*(4), 539–557. https://doi.org/10.1177/0272989X14547740

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press, Oxford (UK). http://dx.doi.org/10.1016/j.futures.2015.07.009

Botega, L. F. D. C., & da Silva, J. C. (2020). An Artificial Intelligence approach to support knowledge management on the selection of creativity and innovation techniques. *Journal of Knowledge Management*, 24(5), 1107-1130. https://doi.org/10.1108/JKM-10-2019-0559

Braun, M., Hummel, P., Beck, S., & Dabrock, P. (2021). Primer on an ethics of AI-based decision support systems in the clinic. *Journal of Medical Ethics*, *47*(12), E3. https://doi.org/10.1136/medethics-2019-105860

Brennan, T. A., Leape, L. L., Laird, N. M., Hebert, L., Localio, A. R., Lawthers, A. G., Newhouse, J. P., Weiler, P. C., & Hiatt, H. H. (1991). Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. *The New England journal of medicine*, *324*(6), 370–376. https://doi.org/10.1056/NEJM199102073240604

Brynjolfsson, E., & McAfee, A. (2014). The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies. *W.W. Norton & Company*.

By, R.T. (2005). Organisational Change Management: A critical review. *Journal of change management*, *5*(4), 369-380.

Camilleri, M. A. (2024). Artificial intelligence governance: ethical considerations and implications for social responsibility. *Expert systems*, 41(7). https://doi.org/10.1111/exsy.13406

Capolupo, P., Messeni Petruzzelli, A., & Ardito, L. (2023). A knowledge-based perspective on transgenerational entrepreneurship: unveiling knowledge dynamics across generations in family firms. *Journal of Knowledge Management*. https://doi.org/10.1108/JKM-05-2023-0451

Caputo, F., Cillo, V., Candelo, E., & Liu, Y. (2019a). Innovating through digital revolution: the role of soft skills and Big Data in increasing firm performance. *Management Decision*, 57(8), 2032-2051.

Caputo, F., Garcia-Perez, A., Cillo, V., & Giacosa, E. (2019b). A knowledge-based view of people and technology: directions for a value co-creation-based learning organisation. *Journal of Knowledge Management*, 23(7), 1314-1334. https://doi.org/10.1108/JKM-10-2018-0645

Caputo, F., Keller, B., Möhring, M., Carrubbo, L., & Schmidt, R. (2023a). Advancing beyond technicism when managing big data in companies' decision-making. *Journal of Knowledge Management*. https://doi.org/10.1108/JKM-10-2022-0794

Caputo, F., Magni, D., Papa, A., & Corsi, C. (2021). Knowledge hiding in socioeconomic settings: matching organizational and environmental antecedents. *Journal of Business Research,* 135, 19-27. https://doi.org/10.1016/j.jbusres.2021.06.012

Caputo, F., Papa, A., Cillo, V., & Del Giudice, M. (2019c). Technology readiness for education 4.0: barriers and opportunities in the digital world. In Opening up education for inclusivity across digital economies and societies (pp. 277-296). *IGI Global*.

Caputo, F., Sepe, F., Di Taranto, E., & Fiano, F. (2023b). Human–technology dichotomy in shaping management history. *Journal of Management History*. https://doi.org/10.1108/JMH-12-2022-0083

Carroll, P.B., Mui, C. (2008). *Billion-dollar lessons: what you can learn from the most inexcusable business failures of the last 25 years*. Penguin Group, New York.

Carter, L., Liu, D., & Cantrell, C. (2020). Exploring the intersection of the digital divide and artificial intelligence: A hermeneutic literature review. *AIS Transactions on Human-Computer Interaction, 12*(4), 253–275. https://doi.org/10.17705/1thci.00138

Castagno, S., & Khalifa, M. (2020). Perceptions of Artificial Intelligence Among Healthcare Staff: A Qualitative Survey Study. *Frontiers in Artificial Intelligence*, *3*. https://doi.org/10.3389/frai.2020.578983

Chen, X., Xie, H., Li, Z., Cheng, G., Leng, M., & Wang, F. L. (2023). Information fusion and Artificial Intelligence for smart healthcare: A bibliometric study. *Information Processing and Management*, *60*(1). https://doi.org/10.1016/j.ipm.2022.103113

Cheng, M., Li, X., & Xu, J. (2022). Promoting Healthcare Workers' Adoption Intention of Artificial-Intelligence-Assisted Diagnosis and Treatment: The Chain Mediation of Social Influence and Human–Computer Trust. *International Journal of Environmental Research and Public Health*, *19*(20). https://doi.org/10.3390/ijerph192013311

Chhabra, B., & Pandey, P. (2023). Job insecurity as a barrier to thriving during COVID-19 pandemic: a moderated mediation model of knowledge hiding and benevolent leadership. *Journal of Knowledge Management*, 27(3), 632-654. https://doi.org/10.1108/JKM-05-2021-0403

Chin, T., Cheng, T.C.E., Wang, C. and Huang, L. (2024). Combining artificial and human intelligence to manage cross-cultural knowledge in humanitarian logistics: a Yin–Yang dialectic systems view of knowledge creation. *Journal of Knowledge Management*, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/JKM-06-2023-0458

Chomutare, T., Tejedor, M., Svenning, T. O., Marco-Ruiz, L., Tayefi, M., Lind, K., Godtliebsen, F., Moen, A., Ismail, L., Makhlysheva, A., & Ngo, P. D. (2022). Artificial Intelligence Implementation in Healthcare: A Theory-Based Scoping Review of Barriers and Facilitators. *International Journal of Environmental Research and Public Health*, *19*(23). https://doi.org/10.3390/ijerph192316359

Choudhury, A. (2022). Factors influencing clinicians' willingness to use an AI-based clinical decision support system. *Frontiers in Digital Health*, *4*. https://doi.org/10.3389/fdgth.2022.920662

Choudhury, A., & Asan, O. (2020). Role of Artificial Intelligence in patient safety outcomes: Systematic literature review. *JMIR Medical Informatics*, *8*(7). https://doi.org/10.2196/18599

Choudhury, A., Asan, O., & Medow, J. E. (2022). Effect of risk, expectancy, and trust on clinicians' intent to use an Artificial Intelligence system – Blood Utilization Calculator. *Applied Ergonomics*, *101*. https://doi.org/10.1016/j.apergo.2022.103708

Chua, I. S., Gaziel-Yablowitz, M., Korach, Z. T., Kehl, K. L., Levitan, N. A., Arriaga, Y. E., Jackson, G. P., Bates, D. W., & Hassett, M. (2021). Artificial intelligence in oncology: Path to implementation. *Cancer medicine*, *10*(12), 4138–4149. https://doi.org/10.1002/cam4.3935

Chui, M., Roberts, R., & Yee, L. (2022). Generative AI is here: How tools like ChatGPT could change your business. *Quantum Black AI by McKinsey*, *20*.

Ciecierski-Holmes, T., Singh, R., Axt, M., Brenner, S., & Barteit, S. (2022). Artificial intelligence for strengthening healthcare systems in low-and middle-income countries: a systematic scoping review. *npj Digital Medicine*, *5*(1), 162. https://doi.org/10.1038/s41746-022-00700-y

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). Science mapping software tools: Review, analysis, and cooperative study among tools. *Journal of the American Society for Information Science and Technology, 62*(7), 1382–1402. https://doi.org/10.1002/asi.21525

Coccia, M. (2016). Radical innovations as drivers of breakthroughs: characteristics and properties of the management of technology leading to superior organisational performance in the discovery process of R&D labs. *Technology Analysis and Strategic Management*, 28(4), 381-395. http://dx.doi.org/10.1080/09537325.2015.1095287

Connelly, C. E., Zweig, D., Webster, J., & Trougakos, J. P. (2012). Knowledge hiding in organizations. *Journal of Organizational Behaviour*, 33(1), 64–88. https://doi.org/10.1002/job.737

Corvello, V., De Carolis, M., Verteramo, S., Steiber, A. (2022). The digital transformation of entrepreneurial work. *International Journal of Entrepreneurial Behaviour and Research*, 28(5), 1167–1183. https://doi.org/10.1108/IJEBR-01-2021-0067

Currie, G., & Hawk, K. E. (2021). Ethical and Legal Challenges of Artificial Intelligence in Nuclear Medicine. *Seminars in nuclear medicine*, *51*(2), 120–125. https://doi.org/10.1053/j.semnuclmed.2020.08.001

D'Este, P., Amara, N., Olmos-Peñuela, J. (2016). Fostering novelty while reducing failure: balancing the twin challenges of product innovation. *Technological Forecasting of Social Change*, 113, 280–292. https://doi.org/10.1016/j.techfore.2015.08.011

Dabić M., Vlačić B., Kiessling T., Caputo A., Pellegrini M. (2021) Serial entrepreneurs: A review of literature and guidance for future research. *Journal of Small Business Management*,1–36. https://doi.org/10.1080/00472778.2021.1969657

Daich Varela, M., Sen, S., De Guimaraes, T. A. C., Kabiri, N., Pontikos, N., Balaskas, K., & Michaelides, M. (2023). Artificial intelligence in retinal disease: Clinical application, challenges, and future directions. *Graefe's Archive for Clinical and Experimental Ophthalmology*. https://doi.org/10.1007/s00417-023-06052-x

Dal Mas, F., Garcia-Perez, A., Sousa, M. J., da Costa, R. L., & Cobianchi, L. (2020). Knowledge translation in the healthcare sector. A structured literature review. *Electronic Journal of Knowledge Management*, 18(3), 198-211. https://doi.org/10.34190/EJKM.18.03.001

Daniel, H., Bornstein, S. S., Kane, G. C., Health and Public Policy Committee of the American College of Physicians, Carney, J. K., Gantzer, H. E., Henry, T. L., Lenchus, J. D., Li, J. M., McCandless, B. M., Nalitt, B. R., Viswanathan, L., Murphy, C. J., Azah, A. M., & Marks, L. (2018). Addressing Social Determinants to Improve Patient Care and Promote Health Equity: An American College of Physicians Position Paper. *Annals of internal medicine*, *168*(8), 577–578. https://doi.org/10.7326/M17-2441

de Hond, A. A. H., Leeuwenberg, A. M., Hooft, L., Kant, I. M. J., Nijman, S. W. J., van Os, H. J. A., Aardoom, J. J., Debray, T. P. A., Schuit, E., van Smeden, M., Reitsma, J. B., Steyerberg, E. W., Chavannes, N. H., & Moons, K. G. M. (2022). Guidelines and quality criteria for Artificial Intelligence-based prediction models in healthcare: A scoping review. *Npj Digital Medicine*, *5*(1). https://doi.org/10.1038/s41746-021-00549-7

De Waal, B., van Outvorst, F., & Ravesteyn, P. (2016). Digital leadership: The objective-subjective dichotomy of technology revisited. In 12th *European Conference on Management, Leadership and Governance* ECMLG 2016 (p. 52).

Del Giudice, M., Scuotto, V., & Papa, A. (2023a). *Knowledge Management and AI in Society 5.0*. Routledge.

Del Giudice, M., Scuotto, V., Orlando, B., & Mustilli, M. (2023b). Toward the human–centered approach. A revised model of individual acceptance of AI. *Human Resource Management Review*, 33(1), 100856. https://doi.org/10.1016/j.hrmr.2021.100856

Del Giudice, M., Scuotto, V., Papa, A., Tarba, S. Y., Bresciani, S., & Warkentin, M. (2021). A self-tuning model for smart manufacturing SMEs: Effects on digital innovation. *Journal of Product Innovation Management*, 38(1), 68-89. https://doi.org/10.1111/jpim.12560

Dewar R. D., Dutton J.E. (1986). The Adoption of Radical and Incremental Innovations: An Empirical Analysis. *Management Science*, 30, 682 695. https://doi.org/10.1287/mnsc.32.11.1422

Donthu N., Kumar S., Mukherjee D., Pandey N., Lim W.M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *J Bus Res*, 133:285–296. https://doi.org/10.1016/j.jbusres.2021.04.070

El Morr, C., & Subercaze, J. (2010). Knowledge management in healthcare. In *Handbook of research on developments in e-health and telemedicine: Technological and social perspectives* (pp. 490-510). IGI Global.

Esmaeilzadeh P., (2020). Use of AI-based tools for healthcare purposes: A survey study from consumers' perspectives. *BMC Medical Informatics and Decision Making*, *20*(1). https://doi.org/10.1186/s12911-020-01191-1

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., ... & Dean, J. (2019). A guide to deep learning in healthcare. *Nature medicine*, *25*(1), 24-29. https://doi.org/10.1038/s41591-018-0316-z

Ettlie J.E., Bridges W.P., O' Keefe R.D. (1984). Organization Strategies and Structural Differences for Radical Versus Incremental Innovation. *Management Science*, 30, 682-695. https://doi.org/10.1287/mnsc.30.6.682

European Commission. (2019). *Ethics guidelines for trustworthy AI*. Publications Office of the European Union. Retrieved November 12, 2024, from https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

European Parliament and Council of the European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation). *Official Journal of the European Union*, L119, 1–88. Retrieved November 12, 2024, from https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng?utm_source=chatgpt.com

European Union. (2024). *Regulation (EU) 2024/1689 of the European Parliament and of the Council of 12 July 2024 laying down harmonized rules on artificial intelligence (Artificial Intelligence Act) and amending certain Union legislative acts*. Official Journal of the European Union, L 257, 1–60. Retrieved November 12, 2024, from https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

Falagas, M. E., Pitsouni, E. I., Malietzis, G. A., & Pappas, G. (2008). Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. *FASEB journal: official publication of the Federation of American Societies for Experimental Biology*, *22*(2), 338–342. https://doi.org/10.1096/fj.07-9492LSF

Felicetti, A.M., Corvello, V. & Ammirato, S. (2024). Digital innovation in entrepreneurial firms: a systematic literature review. *Review of Managerial Science,* 18, 315–362. https://doi.org/10.1007/s11846-023-00638-9

Ferrario, A., Gloeckler, S., & Biller-Andorno, N. (2023). Ethics of the algorithmic prediction of goal of care preferences: From theory to practice. *Journal of Medical Ethics*, *49*(3), 165–174. https://doi.org/10.1136/jme-2022-108371

Franklin, M. I. (2012). Being human and the internet: against dichotomies. *Journal of Information Technology*, 27(4), 315-318.

Freeman C., Perez C., (1986). The Diffusion of Technical Innovation and Changes of Tecno-Economic Paradigm, *Paper Presented at The Conference on Innovation-Diffusion*, Venice, Italy.

Fritsch, S. J., Blankenheim, A., Wahl, A., Hetfeld, P., Maassen, O., Deffge, S., Kunze, J., Rossaint, R., Riedel, M., Marx, G., & Bickenbach, J. (2022). Attitudes and perception of Artificial Intelligence in healthcare: A cross-sectional survey among patients. *Digital Health*, *8*. https://doi.org/10.1177/20552076221116772

Gao, X., & Guan, J. (2009). Networks of scientific journals: An exploration of Chinese patent data. *Scientometrics*, *80*(1), 283-302. https://doi.org/10.1007/s11192-007-2013-4

Gerke, S., Yeung, S., & Cohen, I. G. (2020). Ethical and legal aspects of ambient intelligence in hospitals. *Jama*, *323*(7), 601-602. https://doi.10.1001/jama.2019.21699

Gisbert-Pérez, J., Martí-Vilar, M., & González-Sala, F. (2022). Prospect Theory: A Bibliometric and Systematic Review in the Categories of Psychology in Web of Science. *Healthcare* (Basel, Switzerland), 10(10), 2098. https://doi.org/10.3390/healthcare10102098

Glänzel, W. (2001). National characteristics in international scientific co-authorship relations. *Scientometrics, 51*(1), 69–115. https://doi.org/10.1023/A:1010512628145

Glick, A., Clayton, M., Angelov, N., & Chang, J. (2022). Impact of explainable Artificial Intelligence assistance on clinical decision-making of novice dental clinicians. *JAMIA Open*, *5*(2). https://doi.org/10.1093/jamiaopen/ooac031

Greenes, R. A., Bates, D. W., Kawamoto, K., Middleton, B., Osheroff, J., & Shahar, Y. (2018). Clinical decision support models and frameworks: Seeking to address research issues underlying implementation successes and failures. *Journal of Biomedical Informatics*, *78*, 134–143. https://doi.org/10.1016/j.jbi.2017.12.005

Guo, Y., Hao, Z., Zhao, S., Gong, J., & Yang, F. (2020). Artificial intelligence in health care: Bibliometric analysis. *Journal of Medical Internet Research, 22*(7), e18228. https://doi.org/10.2196/18228

Gupta, S., Kamboj, S., & Bag, S. (2021). Role of Risks in the Development of Responsible Artificial Intelligence in the Digital Healthcare Domain. *Information Systems Frontiers*. https://doi.org/10.1007/s10796-021-10174-0

Gurupur, V., & Wan, T. T. H. (2020). Inherent Bias in Artificial Intelligence-Based Decision Support Systems for Healthcare. *Medicina*, *56*(3), 141. https://doi.org/10.3390/medicina56030141

Harada, T., Shimizu, T., Kaji, Y., Suyama, Y., Matsumoto, T., Kosaka, C., Shimizu, H., Nei, T., & Watanuki, S. (2020). A perspective from a case conference on comparing the diagnostic process: Human diagnostic thinking vs. Artificial intelligence (AI) decision support tools. *International Journal of Environmental Research and Public Health*, *17*(17), 1–6. https://doi.org/10.3390/ijerph17176110

Harari, Y. N. (2019). *Sapiens*. Harper.

Hawkins, R., Blind, K., & Page, R. (Eds.). (2017). *Handbook of innovation and standards*. Edward Elgar Publishing. https://doi.org/10.4337/9781783470082

Heesen, J., Müller-Quade, J., Wrobel, S. *et al*. (2020). Certification of AI systems – Compass for the development and application of trusted AI systems. *White Paper*, Lernende Systeme – Germany's Platform for Artificial Intelligence. https://www.plattform-lernende-systeme.de/publikationen.html

Henderson, R. (1994). Managing innovation in the information age. *Harvard business review*, 72(1), 100-105.

Hillson, D. & Murray-Webster, R. (2007). *Understanding and Managing Risk Attitude* (2nd ed.). Routledge. https://doi.org/10.4324/9781315235448

Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences, 102*(46), 16569–16572. https://doi.org/10.1073/pnas.0507655102

Horgan, D., Romao, M., Morré, S. A., & Kalra, D. (2020). Artificial intelligence: Power for civilization and for better healthcare. *Public Health Genomics*, *22*(5–6), 145–161. https://doi.org/10.1159/000504785

Houfani, D., Slatnia, S., Kazar, O., Saouli, H., & Merizig, A. (2021). Artificial intelligence in healthcare: A review on predicting clinical needs. *International Journal of Healthcare Management*, *15*(3), 267–275. https://doi.org/10.1080/20479700.2021.1886478

Iaia, L., Nespoli, C., Vicentini, F., Pironti, M., & Genovino, C. (2023). Supporting the implementation of AI in business communication: the role of knowledge management. *Journal of Knowledge Management*, (ahead-of-print). https://dx.doi.org/10.1108/JKM-12-2022-0944

Iannuzzi, E., Nigro, C. (2022). *Le sfide dell'innovazione nella Digital Era. Spunti di riflessione sugli aspetti economici, sociali e organizzativi*, Edizioni Scientifiche Italiane, Napoli. ISBN: 9788849550726

Ismail, A. F. M. F., Sam, M. F. M., Bakar, K. A., Ahamat, A., Adam, S., & Qureshi, M. I. (2022). Artificial Intelligence in Healthcare Business Ecosystem: A Bibliometric Study. *International Journal of Online and Biomedical Engineering*, *18*(9), 100–114. https://doi.org/10.3991/ijoe.v18i09.32251

Issac, A. C., Bednall, T. C., Baral, R., Magliocca, P., & Dhir, A. (2022). The effects of expert power and referent power on knowledge sharing and knowledge hiding. *Journal of Knowledge Management*, 27(2), 383-403. https://doi.org/10.1108/JKM-10-2021-0750

Javaid, M., Haleem, A., Singh, R. P., & Suman, R. (2021). Significant applications of big data in Industry 4.0. *Journal of Industrial Integration and Management,* *6*(4). https://doi.org/10.1142/S2424862221500135

Jensen, I., Leith, P., Doyle, R., West, J., Miles, M.P. (2016). Innovation system problems: causal configurations of innovation failure. *Journal of Business Research*, 69(11), 5408–5412. https://doi.org/10.1016/j.jbusres.2016.04.146

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, *2*(4), 230–243. https://doi.org/10.1136/svn-2017-000101

Jin, D., Sergeeva, E., Weng, W.-H., Chauhan, G., & Szolovits, P. (2022). Explainable deep learning in healthcare: A methodological survey from an attribution view. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 14*(4), e1548. https://doi.org/10.1002/wsbm.1548

Johnson, M. P., Zheng, K., Padman, R. (2014). Modeling the udinality of user acceptance of technology with an evidence-adaptive clinical decision support system. *Decision Support Systems*, *57*(1), 444–453. https://doi.org/10.1016/j.dss.2012.10.049

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, *47*(2), 263–291. https://doi.org/10.2307/1914185

Kamoto, S. (2017). Managerial innovation incentives, management buyouts, and shareholders' intolerance of failure. *Journal of Corporate Finance*, 42, 55–74. https://doi.org/10.1016/j.jcorpfin.2016.11.002

Kashyap, S., Morse, K. E., Patel, B., & Shah, N. H. (2021). A survey of extant organizational and computational setups for deploying predictive models in health systems. *Journal of the American Medical Informatics Association*, 28(11), 2445–2450. https://doi.org/10.1093/jamia/ocab154

Kasparick, M., Andersen, B., Franke, S., Rockstroh, M., Golatowski, F., Timmermann, D., Ingenerf, J., & Neumuth, T. (2019). Enabling Artificial Intelligence in high acuity medical environments. *Minimally Invasive Therapy and Allied Technologies*, *28*(2), 120–126. https://doi.org/10.1080/13645706.2019.1599957

Kelley, D. J., Ali, A., & Zahra, S. A. (2013). Where Do Breakthroughs Come From? Characteristics of High-Potential Inventions. *Journal of Product Innovation Management*, 30(6), 1212-1226. https://doi.org/10.1111/jpim.12055

Kerasidou, A. (2020). Artificial intelligence and the ongoing need for empathy, compassion and trust in healthcare. *Bulletin of the World Health Organization*, *98*(4), 245–250. https://doi.org/10.2471/BLT.19.237198 https://doi.org/10.2471/BLT.19.237198

Kettinger, W. J., & Grover, V. (1995). Toward a theory of business process change management. *Journal of management information systems*, *12*(1), 9-30.

Khan, J., Saeed, I., Zada, M., Nisar, H. G., Ali, A., & Zada, S. (2023). The positive side of overqualification: examining perceived overqualification linkage with knowledge sharing and career planning. *Journal of Knowledge Management*, 27(4), 993-1015. https://doi.org/10.1108/JKM-02-2022-0111

Khanijahani, A., Iezadi, S., Dudley, S., Goettler, M., Kroetsch, P., & Wise, J. (2022). Organizational, professional, and patient characteristics associated with artificial intelligence adoption in healthcare: A systematic review. Health Policy and Technology, 11(1), 100602. https://doi.org/10.1016/j.hlpt.2022.100602

Kishor, A., & Chakraborty, C. (2022). Artificial intelligence and internet of things-based healthcare 4.0 monitoring system. Wireless personal communications, 127(2), 1615-1631. https://doi.org/10.1007/s11277-021-08708-5

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and software technology*, *51*(1), 7-15. https://doi.org/10.1016/j.infsof.2008.09.009

Korherr, P., Kanbach, D. (2023). Human-related capabilities in big data analytics: a taxonomy of human factors with impact on firm performance. *Review of Managerial Science*, 17, 1943–1970. https://doi.org/10.1007/s11846-021-00506-4

Kovacs, G., & Croskerry, P. (1999). Clinical decision making: an emergency medicine perspective. *Academic emergency medicine: official journal of the Society for Academic Emergency Medicine*, *6*(9), 947–952. https://doi.org/10.1111/j.1553-2712.1999.tb01246.x

Kraus, S., Schiavone, F., Pluzhnikova, A., & Invernizzi, A. C. (2022). Digital transformation in healthcare: Analyzing the current state-of-research. *Journal of Business Research, 123*, 557–567. https://doi.org/10.1016/j.jbusres.2020.10.030

Kulikowski, C. A. (2022). Ethics in the History of Medical Informatics for Decision-Making: Early Challenges to Digital Health Goals. *Yearbook of Medical Informatics*, *31*(1), 317–322. https://doi.org/10.1055/s-0042-1742491

Kulkov, I. (2023). Next-generation business models for Artificial Intelligence start-ups in the healthcare industry. *International Journal of Entrepreneurial Behaviour and Research*, *29*(4), 860–885. https://doi.org/10.1108/IJEBR-04-2021-0304

Kurzweil, R. (2005). The Singularity is Near: When Humans Transcend Biology. *Viking*. https://doi.org/10.1016/j.techfore.2005.12.002

Lambert, S. I., Madi, M., Sopka, S., Lenes, A., Stange, H., Buszello, C.-P., & Stephan, A. (2023). An integrative review on the acceptance of Artificial Intelligence among healthcare professionals in hospitals. *Npj Digital Medicine*, *6*(1). https://doi.org/10.1038/s41746-023-00852-5

Lee, D., & Yoon, S. N. (2021). Application of Artificial Intelligence-Based Technologies in the Healthcare Industry: Opportunities and Challenges. *International journal of environmental research and public health*, *18*(1), 271. https://doi.org/10.3390/ijerph18010271

Liang, H., & Xue, Y. (2022). Save Face or Save Life: Physicians' Dilemma in Using Clinical Decision Support Systems. *Information Systems Research*, *33*(2), 737–758. https://doi.org/10.1287/isre.2021.1082

Liao, G., Li, M., Li, Y., & Yin, J. (2024). How does knowledge hiding play a role in the relationship between leader–member exchange differentiation and employee creativity? A cross-level model. *Journal of Knowledge Management*, 28(1), 69-84. https://doi.org/10.1108/jkm-01-2023-0046

Long, J., Liu, H., & Shen, Z. (2024). Narcissistic rivalry and admiration and knowledge hiding: mediating roles of emotional exhaustion and interpersonal trust. *Journal of Knowledge Management*, 28(1), 1-26. https://doi.org/10.1108/JKM-11-2022-0860

Loomes, G., & Pogrebna, G. (2014). Testing for independence while allowing for probabilistic choice. *Journal of Risk and Uncertainty*, 49(3), 189–211. http://www.jstor.org/stable/43550243

Lundberg, S. M., Nair, B., Vavilala, M. S., Horibe, M., Eisses, M. J., Adams, T., Liston, D. E., Low, D. K., Newman, S. F., Kim, J., & Lee, S. I. (2018). Explainable machine-learning predictions for the prevention of hypoxaemia during surgery. *Nature biomedical engineering*, 2(10), 749–760. https://doi.org/10.1038/s41551-018-0304-0

Lysaght, T., Lim, H. Y., Xafis, V., & Ngiam, K. Y. (2019). AI-Assisted Decision-making in Healthcare: The Application of an Ethics Framework for Big Data in Health and Research. *Asian Bioethics Review*, *11*(3), 299–314. https://doi.org/10.1007/s41649-019-00096-0

Magni, D., Papa, A., Scuotto, V., & Del Giudice, M. (2023a). Internationalized knowledge-intensive business service (KIBS) for servitization: a microfoundation perspective. *International Marketing Review*. https://doi.org/10.1108/imr-12-2021-0366

Magni, D., Del Gaudio, G., Papa, A. and Della Corte, V. (2023b), Digital humanism and Artificial Intelligence: the role of emotions beyond the human–machine interaction in Society 5.0. *Journal*

of Management History. Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/JMH-12-2022-0084

Magrabi, F., Ammenwerth, E., McNair, J. B., De Keizer, N. F., Hyppönen, H., Nykänen, P., Rigby, M., Scott, P. J., Vehko, T., Wong, Z. S.-Y., & Georgiou, A. (2019). Artificial Intelligence in Clinical Decision Support: Challenges for Evaluating AI and Practical Implications. *Yearbook of Medical Informatics*, *28*(1), 128–134. https://doi.org/10.1055/s-0039-1677903

Mahadevaiah, G., Prasad, R. V., Bermejo, I., Jaffray, D., Dekker, A., & Wee, L. (2020). Artificial intelligence-based clinical decision support in modern medical physics: Selection, acceptance, commissioning, and quality assurance. *Medical Physics*, *47*(5), e228–e235. https://doi.org/10.1002/mp.13562

Malak, J. S., Zeraati, H., Nayeri, F. S., Safdari, R., & Shahraki, A. D. (2019). Neonatal intensive care decision support systems using Artificial Intelligence techniques: A systematic review. *Artificial Intelligence Review*, *52*(4), 2685–2704. https://doi.org/10.1007/s10462-018-9635-1

Maslach, D. (2015). Change and persistence with failed technological innovation. *Strategic Management Journal*, 37(4), 714–723. https://doi.org/10.1002/smj.2358

Massaro, M., Dumay, J., & Guthrie, J. (2016). On the shoulders of giants: undertaking a structured literature review in accounting. *Accounting, Auditing & Accountability Journal*, *29*(5), 767-801.

McCain, K. W. (1991). Mapping economics through the journal literature: An experiment in journal cocitation analysis. *Journal of the American Society for Information Science, 42*(4), 290–296. https://doi.org/10.1002/(SICI)1097-4571(199105)42:4<290::AID-ASI8>3.0.CO;2-9

McCoy, L. G., Nagaraj, S., Morgado, F., Harish, V., Das, S., & Celi, L. A. (2020). What do medical students actually need to know about artificial intelligence?. *NPJ digital medicine*, *3*, 86. https://doi.org/10.1038/s41746-020-0294-7

McLaughlin, P., Bessant, J., Smart, P. (2008). Developing an Organisation Culture to Facilitate Radical Innovation. *International Journal of Technology Management*, 44 (3e4), 298e323. https://doi.org/10.1504/IJTM.2008.021041

McLennan, S., Meyer, A., Schreyer, K., & Buyx, A. (2022). German medical students' views regarding artificial intelligence in medicine: A cross-sectional survey. *PLOS Digital Health, 1*(10), e0000114. https://doi.org/10.1371/journal.pdig.0000114

Mehta, N., Pandit, A., & Shukla, S. (2019). Transforming healthcare with big data analytics and Artificial Intelligence: A systematic mapping study. *Journal of Biomedical Informatics*, *100*. https://doi.org/10.1016/j.jbi.2019.103311

Mele, G., Capaldo, G., Secundo, G., & Corvello, V. (2024), Revisiting the idea of knowledge-based dynamic capabilities for digital transformation, *Journal of Knowledge Management*, Vol. 28 No. 2, pp. 532-563. https://doi.org/10.1108/JKM-02-2023-0121

Meskó, B., Hetényi, G., & Gyorffy, Z. (2018). Will Artificial Intelligence solve the human resource crisis in healthcare? *BMC Health Services Research*, *18*(1). https://doi.org/10.1186/s12913-018-3359-4

Messeni Petruzzelli, A., Albino, V., Carbonara, N., & Rotolo, D. (2010). Leveraging learning behaviour and network structure to improve knowledge gatekeepers' performance. *Journal of Knowledge Management*, 14(5), 635-658. https://doi.org/10.1108/13673271011074818

Messeni Petruzzelli, A., Murgia, G., & Parmentola, A. (2023). Opening the Black Box of Artificial Intelligence Technologies: Unveiling the Influence Exerted by Type of Organisations and Collaborative Dynamics. *Industry and Innovation*, 30(9), 1213-1243. https://doi.org/10.1080/13662716.2023.2213182

Middleton, B., Sittig, D. F., & Wright, A. (2016). Clinical Decision Support: A 25 Year Retrospective and a 25 Year Vision. *Yearbook of Medical Informatics*, S103–S116. https://doi.org/10.15265/IYS-2016-s034

Miller, S., Gilbert, S., Virani, V., & Wicks, P. (2020). Patients' Utilization and Perception of an Artificial Intelligence-Based Symptom Assessment and Advice Technology in a British Primary Care Waiting Room: Exploratory Pilot Study. *JMIR Human Factors*, *7*(3). https://doi.org/10.2196/19713

Moazemi, S., Vahdati, S., Li, J., Kalkhoff, S., Castano, L. J. V., Dewitz, B., Bibo, R., Sabouniaghdam, P., Tootooni, M. S., Bundschuh, R. A., Lichtenberg, A., Aubin, H., & Schmid, F. (2023). Artificial intelligence for clinical decision support for monitoring patients in cardiovascular ICUs: A systematic review. *Frontiers in Medicine*, *10*. https://doi.org/10.3389/fmed.2023.1109411

Mudgal, S. K., Agarwal, R., Chaturvedi, J., Gaur, R., & Ranjan, N. (2022). Real-world application, challenges and implication of Artificial Intelligence in healthcare: An essay. *Pan African Medical Journal*, *43*. https://doi.org/10.11604/pamj.2022.43.3.33384

Murphy, K., Di Ruggiero, E., Upshur, R., Willison, D. J., Malhotra, N., Cai, J. C., Malhotra, N., Lui, V., & Gibson, J. (2021). Artificial intelligence for good health: A scoping review of the ethics literature. *BMC Medical Ethics*, *22*(1). https://doi.org/10.1186/s12910-021-00577-8

Naik, N., Hameed, B. M. Z., Shetty, D. K., Swain, D., Shah, M., Paul, R., Aggarwal, K., Brahim, S., Patil, V., Smriti, K., Shetty, S., Rai, B. P., Chlosta, P., & Somani, B. K. (2022). Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility? *Frontiers in Surgery*, *9*. https://doi.org/10.3389/fsurg.2022.862322

Naiseh, M., Al-Thani, D., Jiang, N., & Ali, R. (2023). How the different explanation classes impact trust calibration: The case of clinical decision support systems. *International Journal of Human Computer Studies*, *169*. https://doi.org/10.1016/j.ijhcs.2022.102941

Nazar, M., Alam, M. M., Yafi, E., & Su'Ud, M. M. (2021). A Systematic Review of Human-Computer Interaction and Explainable Artificial Intelligence in Healthcare with Artificial Intelligence Techniques. *IEEE Access*, *9*, 153316–153348. https://doi.org/10.1109/ACCESS.2021.3127881

Nguyen, M., Pontes, N., Malik, A., Gupta, J., & Gugnani, R. (2024). Impact of high involvement work systems in shaping power, knowledge sharing, rewards and knowledge perception of employees. *Journal of Knowledge Management*. Advance online publication. https://doi.org/10.1108/JKM-04-2023-0345

Nonaka, I. (1991) The Knowledge Creating Company. Harvard Business Review, 69, 96-104.

O'Sullivan, D., Fraccaro, P., Carson, E., & Weller, P. (2014). Decision time for clinical decision support systems. *Clinical medicine*, *14*(4), 338-341. https://doi.org/10.7861/clinmedicine.14-4-338

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Oldenburg, B., Parcel, G. (2008). Diffusion of Health Promotion and Health Education Innovations. In K. Glanz, B. K. Rimer, and F. M. Lewis (eds.), *Health Behaviour and Health Education: Theory, Research, and Practice*. (3rd ed.) San Francisco: Jossey-Bass. ISBN 0470432489, 9780470432488

Orlandi, M. A., Landers, C., Weston, R., & Haley, N. (1990). *Diffusion of health promotion innovations*. In K. Glanz, F. M. Lewis, & B. K. Rimer (Eds.), *Health behaviour and health education: Theory, research and practice* (pp. 288-313). San Francisco: Jossey-Bass. ISBN 978-0-7879-9614-7

Papa, A., Mital, M., Pisano, P., & Del Giudice, M. (2020). E-health and wellbeing monitoring using smart healthcare devices: An empirical investigation. *Technological Forecasting and Social Change*, 153, 119226. https://doi.org/10.1016/j.techfore.2018.02.018

Pee, L. G., Pan, S. L., & Cui, L. (2019). Artificial intelligence in healthcare robots: A social informatics study of knowledge embodiment. *Journal of the Association for Information Science and Technology*, *70*(4), 351–369. https://doi.org/10.1002/asi.24145

Petersson, L., Larsson, I., Nygren, J. M., Nilsen, P., Neher, M., Reed, J. E., Tyskbo, D., & Svedberg, P. (2022). Challenges to implementing Artificial Intelligence in healthcare: A qualitative interview

study with healthcare leaders in Sweden. *BMC Health Services Research*, *22*(1). https://doi.org/10.1186/s12913-022-08215-8

Prakash, S., Balaji, J. N., Joshi, A., & Surapaneni, K. M. (2022). Ethical Conundrums in the Application of Artificial Intelligence (AI) in Healthcare – A Scoping Review of Reviews. *Journal of Personalized Medicine*, *12*(11). https://doi.org/10.3390/jpm12111914

Raja, B. S., & Asghar, S. (2020). Beyond the horizon: A meticulous analysis of clinical decision-making practices. *International Journal of Advanced Computer Science and Applications*, *2*, 691–702. https://doi.org/10.14569/ijacsa.2020.0110287

Rajkomar, A., Dean, J., & Kohane, I. (2019). Machine Learning in Medicine. *The New England journal of medicine*, 380(14), 1347–1358. https://doi.org/10.1056/NEJMra1814259

Ramgopal, S., Sanchez-Pinto, L. N., Horvat, C. M., et al. (2023). Artificial intelligence-based clinical decision support in pediatrics. *Pediatric Research, 93*(2), 334–341. https://doi.org/10.1038/s41390-022-02226-1

Raymond, L., Castonguay, A., Doyon, O., & Paré, G. (2022). Nurse practitioners' involvement and experience with AI-based health technologies: A systematic review. *Applied Nursing Research*, *66*. https://doi.org/10.1016/j.apnr.2022.151604

Rhaiem, K., & Amara, N. (2021). Learning from innovation failures: A systematic review of the literature and research agenda. *Review of Managerial Science, 15*(1), 189–234. https://doi.org/10.1007/s11846-019-00339-2

Richardson, J. P., Curtis, S., Smith, C., Pacyna, J., Zhu, X., Barry, B., & Sharp, R. R. (2022). A framework for examining patient attitudes regarding applications of Artificial Intelligence in healthcare. *Digital Health*, *8*. https://doi.org/10.1177/20552076221089084

Richardson, J. P., Smith, C., Curtis, S., Watson, S., Zhu, X., Barry, B., & Sharp, R. R. (2021). Patient apprehensions about the use of Artificial Intelligence in healthcare. *Npj Digital Medicine*, *4*(1). https://doi.org/10.1038/s41746-021-00509-1

Rogers, E. M., Singhal, A., & Quinlan, M. M. (2014). Diffusion of innovations. In *An integrated approach to communication theory and research* (pp. 432-448). Routledge. ISBN: 9780203887011

Rogers, W. A., Draper, H., & Carter, S. M. (2021). Evaluation of Artificial Intelligence clinical applications: Detailed case analyses show value of healthcare ethics approach in identifying patient care issues. *Bioethics*, *35*(7), 623–633. https://doi.org/10.1111/bioe.12885

Rojas, J. C., Teran, M., & Umscheid, C. A. (2023). Clinician Trust in Artificial Intelligence: What is Known and How Trust Can Be Facilitated. *Critical Care Clinics*, *39*(4), 769–782. https://doi.org/10.1016/j.ccc.2023.02.004

Rothman, A. J., & Salovey, P. (1997). Shaping perceptions to motivate healthy behaviour: The role of message framing. *Psychological Bulletin, 121*(1), 3–19. https://doi.org/10.1037/0033-2909.121.1.3

Rundo, L., Pirrone, R., Vitabile, S., Sala, E., & Gambino, O. (2020). Recent advances of HCI in decision-making tasks for optimized clinical workflows and precision medicine. *Journal of Biomedical Informatics*, *108*. https://doi.org/10.1016/j.jbi.2020.103479

Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach*. Pearson. Upper Saddle River, New Jersey. ISBN-13: 978-0-13-604259-4.

Saheb, T., Saheb, T., & Carpenter, D. O. (2021). Mapping research strands of ethics of artificial intelligence in healthcare: A bibliometric and content analysis. *Computers in biology and medicine*, *135*, 104660. https://doi.org/10.1016/j.compbiomed.2021.104660

Saviano, M., Del Prete, M., Mueller, J., & Caputo, F. (2023). The challenging meet between human and artificial knowledge. A systems-based view of its influences on firms-customers interaction. *Journal of Knowledge Management*, 27(11), 101-111.

Scherer, M. U. (2015). Regulating Artificial Intelligence systems: Risks, challenges, competencies, and strategies. *Harvard Journal of Law & Technology*, 29, 353. http://dx.doi.org/10.2139/ssrn.2609777

Schumpeter, J. A., & Swedberg, R. (1934). *The theory of economic development*. Harvard Economic Studies. https://doi.org/10.4324/9781315135564 (ed. 1983)

Secinaro, S., Calandra, D., Secinaro, A., Muthurangu, V., & Biancone, P. (2021). The role of Artificial Intelligence in healthcare: A structured literature review. *BMC Medical Informatics and Decision Making*, *21*(1). https://doi.org/10.1186/s12911-021-01488-9

Senbekov, M., Saliev, T., Bukeyeva, Z., Almabayeva, A., Zhanaliyeva, M., Aitenova, N., Toishibekov, Y., & Fakhradiyev, I. (2020). The recent progress and applications of digital technologies in healthcare: A review. *International Journal of Telemedicine and Applications*, *2020*. https://doi.org/10.1155/2020/8830200

Serenko, A. (2024). The human capital management perspective on quiet quitting: recommendations for employees, managers, and national policymakers. *Journal of Knowledge Management*, 28(1), 27-43. https://doi.org/10.1108/JKM-10-2022-0792

Shahmoradi, L., Safadari, R., & Jimma, W. (2017). Knowledge Management Implementation and the Tools Utilized in Healthcare for Evidence-Based Decision Making: A Systematic Review. *Ethiopian journal of health sciences*, *27*(5), 541–558. https://doi.org/10.4314/ejhs.v27i5.13

Sharma, A., Thomas, D., Konsynski, B. (2017). Finding the "radicalness" in radical innovation adoption. *Journal of Information Systems Applied Research*, 10(2), 12–20.

Shelmerdine, S. C., Arthurs, O. J., Denniston, A., & Sebire, N. J. (2021). Review of study reporting guidelines for clinical studies using Artificial Intelligence in healthcare. *BMJ Health and Care Informatics*, *28*(1). https://doi.org/10.1136/bmjhci-2021-100385

Shen, J., Zhang, C. J. P., Jiang, B., Chen, J., Song, J., Liu, Z., He, Z., Wong, S. Y., Fang, P.-H., & Ming, W.-K. (2019). Artificial intelligence versus clinicians in disease diagnosis: Systematic review. *JMIR Medical Informatics*, *7*(3). https://doi.org/10.2196/10010

Shinners, L., Aggar, C., Grace, S., & Smith, S. (2020). Exploring healthcare professionals' understanding and experiences of Artificial Intelligence technology use in the delivery of healthcare: An integrative review. *Health Informatics Journal*, *26*(2), 1225–1236. https://doi.org/10.1177/1460458219874641

Shinners, L., Grace, S., Smith, S., Stephens, A., & Aggar, C. (2022). Exploring healthcare professionals' perceptions of Artificial Intelligence: Piloting the Shinners Artificial Intelligence Perception tool. *Digital Health*, *8*. https://doi.org/10.1177/20552076221078110

Siala, H., & Wang, Y. (2022). SHIFTing artificial intelligence to be responsible in healthcare: A systematic review. *Social science & medicine (1982)*, *296*, 114782. https://doi.org/10.1016/j.socscimed.2022.114782

Small, H. G., & Koenig, M. E. D. (1977). *Journal clustering using a bibliographic coupling method. Information Processing & Management, 13(5), 277–288.* https://doi.org/10.1016/0306-4573(77)90017-6

Smallman, S. C. (2022). *An Introduction to International and Global Studies* (3rd ed.). University of North Carolina Press.

Souza-Pereira, L., Ouhbi, S., & Pombo, N. (2021). A process model for quality in use evaluation of clinical decision support systems. *Journal of Biomedical Informatics*, *123*. https://doi.org/10.1016/j.jbi.2021.103917

Sreen, N., Sharma, V., Alshibani, S. M., Walsh, S., & Russo, G. (2024). Knowledge acquisition from innovation failures: a study of micro, small and medium enterprises (MSMEs). *Journal of Knowledge Management*, *28*(4), 947-970. https://doi.org/10.1108/JKM-03-2023-0184

Srivani, M., Murugappan, A., & Mala, T. (2023). Cognitive computing technological trends and future research directions in healthcare – A systematic literature review. *Artificial Intelligence in Medicine*, *138*. https://doi.org/10.1016/j.artmed.2023.102513

Sun, T. Q. (2021). Adopting Artificial Intelligence in public healthcare: The effect of social power and learning algorithms. *International Journal of Environmental Research and Public Health*, *18*(23). https://doi.org/10.3390/ijerph182312682

Sunarti, S., Fadzlul Rahman, F., Naufal, M., Risky, M., Febriyanto, K., & Masnina, R. (2021). Artificial intelligence in healthcare: Opportunities and risk for future. *Gaceta Sanitaria*, *35*, S67–S70. https://doi.org/10.1016/j.gaceta.2020.12.019

Tabassi, E. (2023), Artificial Intelligence Risk Management Framework (AI RMF1.0), *NIST Trustworthy and Responsible AI*, National Institute of Standards and Technology, Gaithersburg, MD. https://doi.org/10.6028/NIST.AI.100-1

Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society*, 40(2), 137–157. https://doi.org/10.1080/14494035.2021.1928377

Takahashi, T., & Vandenbrink, D. (2004). Formative knowledge: from knowledge dichotomy to knowledge geography–knowledge management transformed by the ubiquitous information society. *Journal of Knowledge Management*, 8(1), 64-76. https://doi.org/10.1108/13673270410523916

Teece, D. J. (2023). The evolution of the dynamic capabilities framework. In R. Adams, D. Grichnik, A. Pundziene, & C. Volkmann (Eds.), *Artificiality and sustainability in entrepreneurship*, 113-129. Springer. https://doi.org/10.1007/978-3-031-11371-0_6

Tegmark, M. (2017). *Life 3.0: Being Human in the Age of Artificial Intelligence*. Vintage Books, Knopf, New York (USA). ISBN 978-1-101-94659-6

Topol E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine*, *25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Toscani, G. (2023). The effects of the COVID-19 pandemic for artificial intelligence practitioners: the decrease in tacit knowledge sharing. *Journal of Knowledge Management*, 27(7), 1871-1888. https://doi.org/10.1108/JKM-07-2022-0574

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., Moher, D., Peters, M. D., Horsley, T., Weeks, L., Hempel, S., Akl, E. A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M. G., Garritty, C., ... & Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine, 169*(7), 467–473. https://doi.org/10.7326/M18-0850

Truong, B. T. T., Nguyen, P. V., Vrontis, D., & Ahmed, Z. U. (2024). Unleashing corporate potential: the interplay of intellectual capital, knowledge management, and environmental compliance in enhancing innovation and performance. *Journal of Knowledge Management*, *28*(4), 1054-1073. https://doi.org/10.1108/JKM-05-2023-0389

Tushman, M. L., Anderson P. (1986). Technological discontinuities and organizational environments. *Administrative Science Quarterly*, 31, 439-465. https://doi.org/10.2307/2392832

U.S. Food and Drug Administration. (2018). *Software as a medical device (SaMD): Clinical evaluation*. Retrieved November 13, 2024, from https://www.fda.gov/medical-devices/digital-health-center-excellence/software-medical-device-samd

Ulloa, M., Rothrock, B., Ahmad, F. S., & Jacobs, M. (2022). Invisible clinical labor driving the successful integration of AI in healthcare. *Frontiers in Computer Science*, *4*. https://doi.org/10.3389/fcomp.2022.1045704

Valente, F., Paredes, S., Henriques, J., Rocha, T., de Carvalho, P., & Morais, J. (2022). Interpretability, personalization and reliability of a Machine Learning based clinical decision support system. *Data Mining and Knowledge Discovery*, *36*(3), 1140–1173. https://doi.org/10.1007/s10618-022-00821-8

van de Wetering, R., & Versendaal, J. (2018). How a flexible collaboration infrastructure impacts healthcare information exchange. *BLED 2018 Proceedings*, 12. Retrieved from https://aisel.aisnet.org/bled2018/12

Wadden, J. J. (2021). Defining the undefinable: The black box problem in healthcare Artificial Intelligence. *Journal of Medical Ethics*, *48*(10), 764–768. https://doi.org/10.1136/medethics-2021-107529

Wani, S. U. D., Khan, N. A., Thakur, G., Gautam, S. P., Ali, M., Alam, P., Alshehri, S., Ghoneim, M. M., & Shakeel, F. (2022). Utilization of Artificial Intelligence in Disease Prevention: Diagnosis, Treatment, and Implications for the Healthcare Workforce. *Healthcare (Switzerland)*, *10*(4). https://doi.org/10.3390/healthcare10040608

Westerbeek, L., Ploegmakers, K. J., de Bruijn, G.-J., Linn, A. J., van Weert, J. C. M., Daams, J. G., van der Velde, N., van Weert, H. C., Abu-Hanna, A., & Medlock, S. (2021). Barriers and facilitators influencing medication-related CDSS acceptance according to clinicians: A systematic review. *International Journal of Medical Informatics, 152*, 104506. https://doi.org/10.1016/j.ijmedinf.2021.104506

White, H. D., & Griffith, B. C. (1981). Author co-citation: A literature measure of intellectual structure. *Journal of the American Society for Information Science, 32*(3), 163–171. https://doi.org/10.1002/asi.4630320302

White, H. D., & McCain, K. W. (1998). Visualizing a discipline: An author cocitation analysis of information science, 1972–1995. *Journal of the American Society for Information Science, 49*(4), 327–355. https://doi.org/10.1002/(SICI)1097-4571(19980401)49:4<327::AID-ASI4>3.0.CO;2-4

Wilson, A., Saeed, H., Pringle, C., Eleftheriou, I., Bromiley, P. A., & Brass, A. (2021). Artificial intelligence projects in healthcare: 10 practical tips for success in a clinical environment. *BMJ Health and Care Informatics*, *28*(1). https://doi.org/10.1136/bmjhci-2021-100323

Wolff, J., Pauling, J., Keck, A., & Baumbach, J. (2021). Success Factors of Artificial Intelligence Implementation in Healthcare. *Frontiers in Digital Health*, *3*. https://doi.org/10.3389/fdgth.2021.594971

Wu, Z., Zhou, X., Wang, Q., & Liu, J. (2023). How perceived overqualification influences knowledge hiding from the relational perspective: the moderating role of perceived overqualification differentiation. *Journal of Knowledge Management*, 27(6), 1720-1739. https://doi.org/10.1108/JKM-04-2022-0286

Xu, Q., Xie, W., Liao, B., Hu, C., Qin, L., Yang, Z., Xiong, H., Lyu, Y., Zhou, Y., & Luo, A. (2023). Interpretability of Clinical Decision Support Systems Based on Artificial Intelligence from Technological and Medical Perspective: A Systematic Review. *Journal of Healthcare Engineering*, *2023*. https://doi.org/10.1155/2023/9919269

Yan, E., & Ding, Y. (2012). Scholarly network similarity: How bibliographic coupling networks, citation networks, cocitation networks, topical networks, coauthorship networks, and coword networks relate to each other. *Journal of the American Society for Information Science and Technology, 63*(7), 1313–1326. https://doi.org/10.1002/asi.22680

Yang, C. C. (2022). Explainable Artificial Intelligence for Predictive Modeling in Healthcare. *Journal of Healthcare Informatics Research*, *6*(2), 228–239. https://doi.org/10.1007/s41666-022-00114-1

Yoon, S. N., & Lee, D. (2019). Artificial intelligence and robots in healthcare: What are the success factors for technology-based service encounters? *International Journal of Healthcare Management*, *12*(3), 218–225. https://doi.org/10.1080/20479700.2018.1498220

Young, A. T., Amara, D., Bhattacharya, A., & Wei, M. L. (2021). Patient and general public attitudes towards clinical Artificial Intelligence: A mixed methods systematic review. *The Lancet Digital Health*, *3*(9), e599–e611. https://doi.org/10.1016/S2589-7500(21)00132-1

Zhang, J., & Zhang, Z. M. (2023). Ethics and governance of trustworthy medical artificial intelligence. *BMC medical informatics and decision making*, 23(1), 7. https://doi.org/10.1186/s12911-023-02103-9

Zhao, D., & Strotmann, A. (2008). Information science during the first decade of the web: An enriched author co-citation analysis. *Journal of the American Society for Information Science and Technology, 59*(6), 916–937. https://doi.org/10.1002/asi.20799

Zupic, I., and Čater, T. (2015). Bibliometric Methods in Management and Organization. *Organizational Research Methods*, 18, 429-472. https://doi.org/10.1177/1094428114562629

Chapter 2. **Navigating AI Governance in Healthcare: Insights from the European Commission's Official Communications**

Simona Curiello[1, *], Enrica Iannuzzi[2], Claudio Nigro[2]

[1] Department of Social Sciences, University of Foggia, Italy

[2] *Department of Economics, University of Foggia, Italy*

[*] Corresponding Author. Email: simona.curiello@unifg.it

Abstract

The advent of Artificial Intelligence (AI) has the potential to drive a rapid transformation across various sectors, including healthcare, with the capacity to revolutionize processes, decision-making, and service delivery. However, this paradigm shift is accompanied by a concomitant series of critical regulatory and ethical challenges. Relying on the pillars of the New Institutional Theory, this study undertakes a comprehensive examination of the European Commission's formal communications on AI in healthcare, with the aim of identifying: the principal themes; the potential alignment of AI strategies with European Identity politics; the policy recommendations provided by the European Union (EU); and the regulatory challenges and ethical concerns. A multi-method approach was employed, integrating both computational text analysis techniques and a systematic content analysis of 95 documents published between 2018 and 2025. The analysis revealed that the Commission's discourse places significant emphasis on themes such as innovation, public trust, risk management, and the ethical integration of AI. This study uniquely contributes to the discourse on AI governance by demonstrating how institutional communication shapes policy and reinforces Europe's commitment to "*human-centric*" and "*trustworthy AI*", which is firmly embedded within the broader digital sovereignty agenda of the European Union and the regulatory frameworks it has established, most notably the Artificial Intelligence Act. The findings provide theoretical insights into institutional discourse and offer practical implications for policymakers and healthcare stakeholders navigating AI regulations.

Keywords: Artificial Intelligence (AI), Healthcare, Ethics, Regulatory Frameworks

Introduction

The European Commission (2018) defines Artificial Intelligence (AI) as "*systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals*"[3].

The accelerated evolution of AI technologies has outstripped the development of corresponding regulatory and normative frameworks, particularly in domains that are pivotal to the preservation of democratic values and human rights (Coeckelbergh, 2020; Dignum, 2019; Marcus & Davis, 2019 Ulnicane *et al*., 2021; Floridi, 2022; Wörsdörfer, 2023). While governments and international institutions endeavour to address this deficit, an increasing number of nations have introduced

[3] European Commission. (2018). *Communication artificial intelligence for Europe.* Policy and legislation. Retrieved from https://digital-strategy.ec.europa.eu/en/library/communication-artificial-intelligence-europe

national AI strategies and policies to facilitate the responsible utilisation of these technologies (Taeihagh, 2021).

At present, over 175 countries, companies, and organisations have published documents that set out ethical guidelines for AI, which reflects a heightened focus on regulating and safeguarding these systems. However, the efforts to regulate AI have been somewhat fragmented, with multiple, often competing, initiatives. Despite this fragmentation, indications of convergence are emerging, particularly as the United States, Europe, and other nations coalesce around shared principles (Riaz, 2009; Dickinson *et al*., 2021).

Furthermore, while global organisations such as the Organisation for Economic Co-operation and Development (OECD) and the United Nations Educational, Scientific and Cultural Organisation (UNESCO) have established overarching AI principles with the objective of shaping ethical standards and practices, the European Union (EU) stands out due to its comprehensive, legally binding regulatory approach.

In particular, following this approach, the European Commission (EC) has placed an increasing emphasis on the strategic potential of AI in addressing critical societal and human challenges, while also shaping the EU's vision, priorities, and strategy for AI integration (European Commission, 2018). A significant milestone in this effort was the publication of the *Ethics Guidelines for Trustworthy Artificial Intelligence* in 2019[4], which delineated seven fundamental principles for responsible AI development. These criteria – rooted in fundamental rights and ethical principles – prioritize fairness, accountability, and transparency in AI governance. The guidelines align with broader global debates on AI ethics (Floridi *et al*., 2018; Jobin *et al*., 2019), which emphasize the importance of balancing innovation with societal safeguards.

Overall, the EU's AI strategy is indicative of the intersection of societal needs, ethical concerns, and regulatory challenges (Winfield & Jirotka, 2018). The European Union has distinguished itself as a global leader in the governance of artificial intelligence (AI) through the establishment of the *Artificial Intelligence Act* (AIA), a seminal piece of legislation that aims to establish a harmonised regulatory framework for the development and deployment of AI. This initiative represents a significant departure from voluntary ethical guidelines or fragmented national regulations, positioning the EU at the vanguard of AI governance on a global scale.[5] Moreover, the EU's regulatory strategies have been shown to exert a notable influence on global AI policy, a phenomenon known as the *Brussels Effect* (Almada & Radu, 2024). This effect refers to the tendency for European

[4] European Commission. (2019). *Ethics guidelines for trustworthy AI*. Retrieved from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai
[5] European Union. (n.d.). *EU Artificial Intelligence Act*. Retrieved, from https://artificialintelligenceact.eu/

standards to shape international markets and governance models. In light of the EU's pivotal role in AI policymaking and its structured approach to AI ethics and regulation, this study aims to examine the influence of these policies on AI governance and global regulatory trends by exploring the EU's institutional discourse.

To date, scholars have critiqued such high-level ethical frameworks for their lack of enforceability and sector-specific guidance (Hagendorff, 2020). For instance, while the EC promotes principles of equitable AI, the operationalization of these principles in highly regulated and high-stakes domains such as business and healthcare remains a challenge, with scholars calling for clearer regulatory mechanisms to translate these guidelines into practice (Mittelstadt, 2019; Morley *et al*., 2021).

As previously stated, there remains a critical gap in understanding how institutional discourse shapes AI governance. Existing studies often focus on regulatory frameworks and journal publications (af Malmborg & Trondal, 2021; Ulnicane *et al*., 2021), yet the role of formal communication documents in influencing policy and public perception is less explored (Birkstedt *et al*., 2023).

To address this gap, our study examines the European Commission's formal communications to identify key themes, strategic alignments, and regulatory challenges.

In particular, in the context of healthcare, a sector characterised by ethical sensitivity, the accelerated development of AI has the potential to engender both transformative prospects and substantial regulatory and ethical dilemmas (Naik *et al*., 2022; Prakash *et al*., 2022).

In contrast to other sectors, errors in algorithmic decision-making in healthcare can have dire consequences, as these decisions can directly impact patient outcomes (Liang & Xue, 2022; Zhang & Zhang, 2023). Moreover, the necessity of holistic patient data, the intricacies of co-morbidities, and the imperative for precise diagnostics create a landscape where companies might be incentivized to navigate regulatory grey areas (Gerke *et al*., 2020; Zhang & Zhang, 2023). The potential for regulatory misinterpretation or circumvention is heightened in healthcare compared to other sectors (Hyman, 2001; Krause, 2002). These dynamics make healthcare a critical domain for studying AI governance communication. Furthermore, as AI-driven solutions increasingly influence diagnostics, treatment planning, and personalized medicine, ensuring compliance with ethical and regulatory standards is paramount (European Commission, 2019; Lysaght *et al*., 2019; Naik *et al*., 2022; Saheb *et al*., 2021).

The intricacies inherent in this domain, coupled with the paucity of studies that delve into the specific influence of institutional discourses on shaping AI policies in healthcare have guided the authors' decision to undertake a focused analysis of the EU's communication on AI, with a specific focus on the healthcare sector.

This study aims to address the following research questions:

RQ1. What are the principal ***themes*** and ***subjects*** addressed in the European Commission's communications on AI in healthcare?

RQ2. How do the European Commission's AI strategies correspond with broader ***European Identity (EI)*** politics?

RQ3. What ***policy recommendations*** are proposed in the European Commission's communications on the integration of AI in healthcare?

By analysing formal communication documents, this study broadens the scope of discourse analysis in organizational research, emphasizing the role of public institutions in AI regulation.

From a practical perspective, the findings of this paper aim to provide actionable insights for regulators and policymakers on how discourse shapes public perception and policy formation around AI in healthcare.

The paper is organised as follows: Section 2 provides a detailed overview of the theoretical framework. Section 3 outlines the research methodology, which includes the qualitative content analysis of European Commission communication documents. Section 4 presents the key findings, categorising the dominant themes in AI governance discourse. Section 5 discusses the implications of these findings for both policy and practice. Finally, the paper concludes with a summary of the contributions, limitations, and directions for future research.

Literature Review

The European Union has placed significant emphasis on the concept of *digital sovereignty* as a cornerstone of its digital and AI policies (Mügge, 2024). This concept is indicative of the EU's aspiration to exercise control over key dimensions of the digital realm, including physical infrastructure, standards and rules, and data and content. As articulated by European Commission President Ursula von der Leyen in 2020, *digital sovereignty*, in conjunction with "*open strategic autonomy*", constitutes a vital component of Europe's broader global ambitions. This approach emphasises the EU's commitment to establish "*responsible AI*", "*human-centric AI*", and "*trustworthy AI*" as fundamental tenets of its regulatory framework, distinguishing it from the commercially-driven models of the United States and the state-led strategies of China.

In the context of AI, various stakeholders, including governments, the private sector, and civil society, are engaged in developing governance models that are shaped by shared principles, norms, rules, decision-making processes, and economic factors. For the purposes of this paper, norms are defined as the "*collective expectations for appropriate behaviour of actors with a particular*

identity"[6]. These norms exist in the "*grey zone*" between law and politics, often referred to as soft law. Unlike formal laws, norms do not impose legally binding obligations but instead set expectations for conduct and allow actors to experiment with certain rules and principles before they might become formalized into law.

The EU's proposed Artificial Intelligence Act (AIA) is at the core of this framework, employing a "risk-based" approach to categorise AI systems. For instance, systems posing "unacceptable risks", such as biometric surveillance technologies, are prohibited outright. Conversely, systems designated as "high-risk", including predictive algorithms in hiring processes and AI-powered medical diagnostic tools, are subject to rigorous pre-market conformity assessments. This ensures that such systems are compliant with safety, ethical, and transparency standards prior to their release into the market. The *European Declaration on Digital Rights and Principles*[7] reinforces this view by stating that technology should unite people and contribute to a fair, inclusive society and economy. The European Commission has similarly articulated a vision for a digital society by 2030[8], emphasizing that no one should be left behind in the digital transformation, with particular attention given to those who may be vulnerable or least likely to benefit.

The EU's pursuit of digital sovereignty, characterised by its endeavours to exercise control over digital infrastructure, standards, and data governance, finds close alignment with the dynamics delineated by *New Institutional Theory*. This theoretical background underscores the manner in which institutional frameworks are established and deconstructed over time, frequently reflecting the actions and strategies of influential actors, known as *Institutional Entrepreneurs* (IEs) (DiMaggio, 1988; Dorado, 2005; Battilana, 2006). These actors engage in what is termed "Institutional Work", actively shaping, modifying, or dismantling institutional norms and practices to align them with their objectives (Leca e Naccache, 2006; Hardy e Maguire; 2008).

In the context of the EU's digital policies, particularly its AIA, IEs comprising policymakers, industry leaders, and advocacy groups play a crucial role. By leveraging legitimacy and strategic influence, these actors drive the development of regulatory norms, including the "risk-based" approach central to AI governance. This framework, anchored in the concept of *digital sovereignty*, gradually takes shape over time, evolving into a set of institutional prescriptions that actors seek to

[6] Katzenstein, P. J. (Ed.). (1996). *The culture of national security: Norms and identity in world politics*. Columbia University Press.
[7] European Commission. (2022). *European Declaration on Digital Rights and Principles*. Retrieved from https://digital-strategy.ec.europa.eu/en/library/european-declaration-digital-rights-and-principles
[8] European Commission. Europe's Digital Decade: digital targets for 2030. Retrieved from https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_en

align with, often in pursuit of legitimacy. This process has been theoretically articulated by Meyer and Rowan (1977) and Powell and DiMaggio (1991).

The interplay between digital sovereignty and the institutional framework becomes evident when considering the EU's emphasis on "human-centric AI" and "trustworthy AI." These principles, it should be noted, represent more than mere regulatory ambitions; they also embody institutionalised norms that have been shaped through a dialectic process of negotiation, decision-making and communication among a diverse range of stakeholders (Hardy & Maguire, 2017).

The insights offered by neo-institutionalist theory are particularly valuable in this regard, as it focuses on key dimensions such as:

- *Institutional Entrepreneurship*: this dimension examines the role of IEs in influencing and reshaping institutional contexts, including their use of institutional strategies to navigate and shape the regulatory landscape (DiMaggio, 1988; Hardy & Maguire, 2008).
- *Institutional Work*: highlighting the broader processes of constructing, maintaining, and transforming institutions through actions by both powerful and less powerful actors, emphasising collaborative dynamics in institutional evolution (Lawrence & Suddaby, 2006; Battilana *et al*., 2009).

A more detailed examination of the theoretical framework of this paper reveals its foundation in the role of communication as a pivotal mechanism in the formation of institutional norms and practices.

From the perspective of *Institutional Work*, communication mechanisms serve as strategic levers, essential for constructing and disseminating discourses that guide decision-making processes leading to constraints and norms (regulatory framework) within a concrete system of action, resulting from a mix of games (conflicts/negotiations), values, and rules that temporarily structure a local order (Clegg *et al*., 2006; Greenwood *et al*., 2008).

It is through communication that institutional actors are able to establish their legitimacy, influence the decision-making process, and create the symbolic and cultural meanings that underpin regulatory frameworks (Suddaby, 2011). Institutional change is often the result of a dialectic process involving competing discourses that shape and reshape norms over time. In the context of AI governance, the creation of stratified texts –such as policy documents, regulatory proposals, and strategic communications– that legitimize and promote the adoption of responsible AI practices is imperative (Maguire & Hardy, 2006).

In the realm of AI adoption across entrepreneurial sectors, IEs leverage various resources:

1. *Reputation Capital*. Success in implementing and innovating AI technologies enhances the credibility of IEs, fostering a positive reputation crucial for influencing stakeholders (Zucker,

1986; Barney, 1991; Nooteboom, 1996; Fombrun, 1996; Ferguson *et al*., 2000; Roberts & Dowling, 2002; Rindova *et al*., 2005).

2. *Information Asymmetries*. Effectively navigating complex AI technologies and interpreting related information gives Entrepreneurs an advantage, reducing information gaps and aligning AI initiatives with stakeholder goals (Aguirre *et al*., 2020; Di Porto & Zuppetta, 2021; Radhakrishnan *et al*., 2022; Prasad Babu & Vasumathi, 2023).
3. *Accreditation Toward Institutions*. Recognition from regulatory bodies, certifications in AI ethics, or partnerships with established AI organizations provide legitimacy, enabling Entrepreneurs to actively participate in policy discussions and shape industry standards (Mead, 1934; Parsons, 1951; Strauss, 1978; Giddens, 1984; Barley & Tolbert, 1997).
4. *Financial Resources*. Crucial for AI adoption, financial resources cover technology acquisition, hiring skilled professionals, and research. Access to substantial funds empowers institutional agents to effectively implement and scale AI initiatives (Riaz *et al*., 2011).
5. *Apical Roles in AI Communication Processes*. Holding influential roles in communication processes allows Entrepreneurs to shape how AI technologies are portrayed to the market, clients, and regulatory bodies. Steering the narrative around the *ethical use*, *benefits*, and *risks* of AI enables Entrepreneurs to influence market trends and regulatory discussions (Brauner *et al*., 2023).

Institutional Work theory emphasizes that institutional activities are often shaped and legitimized through discourse. Communication, whether verbal or written, serves as a fundamental mechanism to influence regulatory frameworks and institutional landscapes. In this regard, the *organizational discourse theory* (Putnam & Fairhurst, 2001) is widely used in institutionalist literature to analyse *Institutional Work* processes (Lawrence & Suddaby, 2006; Hardy, 2011). Discourse theory encapsulates a broad range of methods and approaches to organizational analysis, focusing on "organizational discourse" as a structured set of texts embedded within written and verbal practices. These texts, alongside cultural artifacts, contribute to the construction of concepts, strategies, and policy frameworks as they are produced, disseminated, and consumed (Grant *et al*., 2004; Hardy & Maguire, 2017).

As anticipated before, *New Institutional Theory* emphasizes the role of institutional structures, norms, and discursive practices in shaping governance frameworks and decision-making (DiMaggio & Powell, 1983; Scott, 2008). Institutions, in this sense, serve as both enablers and constraints for policy development, influencing regulatory discourses and standard-setting processes (North, 1990; Greenwood *et al*., 2008).

In the context of this study, Institutional Theory provides a lens through which we can understand how the EC constructs its regulatory and ethical discourse around AI governance in healthcare (Zucker, 1987). Institutional work aligns with the coding taxonomy in Appendix 1 by categorizing discourses that legitimize AI policies and governance mechanisms (Suchman, 1995).

Each category can be mapped onto key dimensions of Institutional Theory:

1. *Descriptive Framework on AI & European Identity (Normative Pressure)*

The EU's emphasis on human-centric and trustworthy AI reflects normative isomorphism (DiMaggio & Powell, 1983). The coding taxonomy captures this through references to the EU's AI Act and European Identity, highlighting the institutional legitimacy of AI regulations (Hall & Taylor, 1996; Scott, 2014).

2. *Regulatory Landscape & Risk Categorization (Regulatory Institutionalization)*

The categorization of AI systems into high-risk, limited-risk, and minimal-risk aligns with the regulatory logic of institutional frameworks (Meyer & Rowan, 1977). Institutional actors, such as the EC, act as institutional entrepreneurs who shape AI governance by embedding risk assessment into the regulatory process (Fligstein, 2001; Maguire *et al*., 2004).

3. *Ethical and Transparency Issues (Mimetic Processes)*

The concern with algorithmic bias, privacy, and ethical considerations in AI governance mirrors institutional processes where policymakers model their frameworks based on perceived best practices and ethical norms (Scott, 1995; Powell & DiMaggio, 1991). The AI governance framework is shaped by institutional isomorphism, wherein global regulatory structures and AI governance trends influence EU policies (Berger & Luckmann, 1966; Czarniawska & Sevon, 1996).

4. *Data Governance and Digital Sovereignty (Institutional Entrepreneurship)*

The focus on data protection, privacy, and interoperability demonstrates the EU's role as an institutional entrepreneur (Battilana *et al*., 2009). The taxonomy reflects institutional theory's insights on how power dynamics and strategic actions shape regulatory norms (Hardy & Maguire, 2008; Hoffman, 1999).

5. *Healthcare Applications and Member States' Digital Level (Institutional Diffusion)*

The taxonomy's inclusion of AI's role in health data governance and digital health adoption across member states reflects institutional diffusion – where policies spread across national and regional levels through regulatory alignment and policy learning (Tolbert & Zucker, 1983; Sahlin & Wedlin, 2008).

Methodology

The methodology employed in this research involved a systematic approach to gather and analyse communications and speeches on AI in healthcare from the European Commission's Press Corner website (https://ec.europa.eu/commission/presscorner/). The analysis of the collected data entailed a content analysis, focusing on the communication surrounding AI in healthcare as articulated by members of the European Commission in the documents published on the official Press Corner website, from 2018 to date. This period is significant as it marks the EU's intensification of AI regulation, including the publication of the AI Act.

Initially, a comprehensive search was conducted using specific keywords related to AI and healthcare, resulting in a pool of 202 search results. Subsequently, the authors downloaded the identified documents and meticulously screened them for eligibility criteria, ultimately selecting 95 papers that focused on the European formal communication on AI and its healthcare applications. To identify these papers, a detailed examination was performed, targeting specific keywords such as "health", "medicine", "pharma", "artificial intelligence", and "AI" within the text. The selected documents were then catalogued in an Excel spreadsheet. To investigate how the European Commission reconciles its AI strategy with European Identity politics in its framing practices, the authors examined four types of EC documents released between 15 March 2018 and 13 March 2025: speeches, press releases, Q&A documents and statements. These documents serve various purposes, from disseminating timely information and detailed explanations to articulating strategic visions and policy responses, collectively shaping the narrative on AI and digital innovation in the EU (Figueira, 2017).

In this study, content analysis was applied as a mixed quali-quantitative approach, combining computational and interpretative techniques to identify and quantify recurring policy themes (Krippendorff, 1980; Patton, 2014; Ulnicane *et al*., 2021). Analytical constructs or rules of inference were used to link the text to research questions (White & Marsh, 2006).

The analysis, spanning July 2024 to March 2025, followed a two-step coding strategy integrating deductive and inductive approaches.

In the first stage, a deductive coding framework was developed in conformity with the New Institutional Theory, using predefined analytical categories such as *regulative mechanisms, normative frameworks,* and *legitimacy narratives*. Deductive category assignment (Mayring, 2014) was applied for the entire data corpus in NVivo®, where the coding units were defined as sentences or logical sentence parts containing one or more of the target keywords (e.g., AI, health, data, ethics). This method is suitable due to the descriptive nature of the primary research question, which seeks to identify the elements present in the data rather than explain them (RQ1: themes and subjects).

Axial coding techniques were used to group related codes into more general categories (Saldaña, 2021). An initial codebook was drafted, classifying keyword mentions into core categories – *AI (Artificial Intelligence)*, *healthcare*, and *medicine* – which served as the primary search and classification anchors. These represented the foundational keywords for document identification. However, additional related terms (e.g., *health data*, *ethics*, *risk*, *digital health*, *trust*) were subsequently incorporated during the coding process to capture emergent concepts. Following this approach, four broad discourses encapsulating how the Commission frames AI – 1) *Ethical and Legal Regulation*, 2) *Innovation and Competitiveness*, 3) *Data Governance and Trust*, and 4) *European Leadership and Identity* – were created by combining categories that shared conceptual overlap. Two authors independently coded and analysed the full selected documents, comparing findings and resolving areas of confusion. Changes were made to classification descriptions (Patton, 2014), incorporated into the codebook, and final coding established.

While the initial search strategy used a combination of AI- and healthcare-related keywords to identify relevant EC communications, the subsequent analyses were conducted on the *full text* of each of the 95 selected documents. This means that frequency counts (e.g., for 'health', 'data', or 'intelligence') reflect the actual distribution of terms across full documents, rather than a search bias or excerpt-level sampling. Thus, the thematic prominence of healthcare-related vocabulary emerged empirically from the corpus as a whole, not from the inclusion criteria alone. Illustrative examples of codes aggregated under each discourse and representative quotations from the coded corpus are provided in *Appendix 1*.

In the second stage, Latent Dirichlet Allocation (LDA) topic modelling was carried out for the entire corpus employing R (v. 4.3.2) to complement and validate the deductive coding. The computational analysis aimed to uncover latent thematic patterns independent of the pre-established codebook. The dataset was first loaded and standardized through a pre-processing phase that included text cleaning, lemmatization, and the extraction of frequent bigrams and trigrams (Jurafsky & Martin, 2021). This approach ensured to preserve key multi-word expressions (e.g., "*artificial intelligence*", "*data protection*") for further analysis (Jurafsky & Martin, 2021).

The processed text has then been transformed into a *Document-Term Matrix* (DTM), which represents the frequency of each term in each document (Boyd-Graber *et al*., 2017). The number of topics was determined by model fit criteria. The final LDA model has been trained on the DTM, uncovering latent thematic structures in the text. LDA assumes that each document is a mixture of topics and each topic is a mixture of words, making it particularly useful for unsupervised thematic extraction in policy communications (Griffiths & Steyvers, 2004).

The results from the LDA were interpreted using a qualitative approach and compared with the codes extracted in NVivo®. Topics identified through LDA informed the refinement of qualitative categories, allowing convergence between *data-driven* clusters and *theory-driven* codes. In this sequential process, the LDA acted as an inductive validation step, confirming and expanding the results of the prior deductive analysis.

By doing so, this process allowed both the computational objectivity and interpretive depth: while NVivo® coding was grounded in theoretical constructs, topic modelling through LDA picked out themes that might otherwise have remained undetected.

Finally, interpretive synthesis of manually coded data was used to address RQ2 and RQ3. For RQ2, themes identified through qualitative coding were conceptually mapped onto dimensions of European Identity politics, reflecting how discursive framings align with broader EU values and principles. For RQ3, the policy recommendations expressed in the communications were analysed in relation to coded categories that corresponded to regulatory mechanisms. This guaranteed methodological consistency for each of the three research questions.

Findings

An examination of the European Commission's communications on AI in healthcare reveals several key themes that align with the core objectives of the study. The analysis integrates results from deductive qualitative coding in NVivo® and inductive topic modelling (LDA), ensuring both interpretive and data-driven robustness.

Dominant Lexical and Thematic Patterns

The results presented in this section derive from the NVivo®-based qualitative coding and lexical frequency analysis of the 95 European Commission communications included in the corpus. The most frequently cited terms serve to emphasise the centrality of healthcare in the general disclosure on AI by the members of the EC.

The term "*health*" appears as the dominant expression (2,282 occurrences), followed by "*data*" (2,032) and "*intelligence*" (1,751). This reflects a pronounced emphasis on the deployment of AI-driven systems within the context of healthcare, with a particular focus on the role of data in these innovations.

The terms "risk" (418 occurrences), "trustworthy" (362), and "rules" (296) were used in a variety of communications, most often in relation to legal frameworks, ethical standards, and technical compliance.

According to a contextual reading of these references:

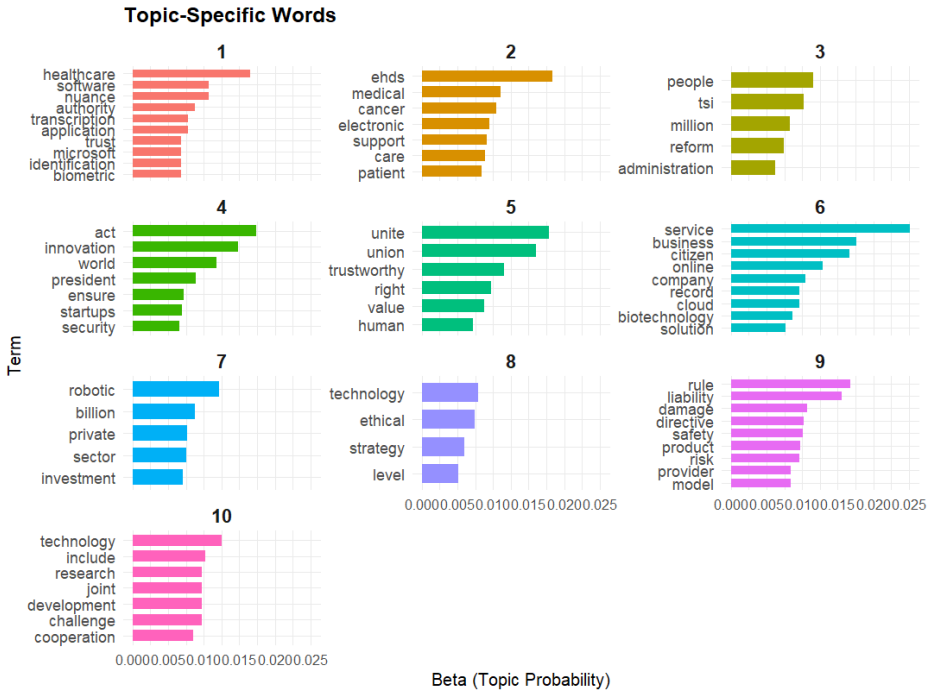
- "trustworthy" was frequently associated with the EU's Trustworthy AI framework and calls for human-centric innovation (*Tech and Geopolitics: Building European Resilience in the Digital Age – Speech by Commissioner Thierry Breton*, 5 September 2023; *President von der Leyen's Speech at the High-Level Opening Session of the 2021 Digital Assembly: "Leading the Digital Decade"*, 1 June 2021);
- "rules" was frequently used to indicate legislative safeguards or standard-setting initiatives (*Commission launches AI innovation package to support Artificial Intelligence startups and SMEs*, Press Release, 24 January 2024; *Europe Fit for the Digital Age: Commission proposes new rules and actions for excellence and trust in Artificial Intelligence*, Press Release, 21 April 2021); and
- "risk" was typically used in relation to data misuse, algorithmic bias, or patient safety (*U.S.-EU Summit Joint Statement*, 20 October 2023; *Questions & Answers: AI Liability Directive*, 28 September 2022; *EU-US Trade and Technology Council Inaugural Joint Statement*, 29 September 2021).

When combined, these contextual applications support the Commission's focus on regulating AI technologies and guaranteeing their security and ethical soundness in delicate fields like healthcare. The term "public" was referenced frequently (247 times), in relation to public interest, public trust and public health. This vocabulary reflects the Commission's view of AI as a technology that serves the public good – a tool that enhances transparency and protects collective welfare – rather than as a purely market-driven innovation.

Integration of Computational and Qualitative Results

The Latent Dirichlet Allocation (LDA) topic modelling analysis, computed in R, systematically identifies clusters of semantically related terms, revealing dominant themes within the European Commission's communications on AI in healthcare. This computational technique, widely recognized in policy analysis and digital text mining (Blei *et al*., 2003; Jacobi *et al*., 2016), enables the detection of latent structures in the dataset, allowing for an empirically grounded identification of discourse trends. **Errore. L'origine riferimento non è stata trovata.** visualises the aggregation of 38 qualitative codes into four higher-order discourses. Each discursive cluster was derived through iterative grouping of semantically related categories identified during coding (see *Appendix 1* for examples of underlying codes).

Figure 1 – Thematic extraction process through computational LDA



Source: our elaboration

To further validate the emerging themes, the results of the LDA topic modelling were cross-referenced with findings from the deductive qualitative coding conducted in NVivo®. This integration enabled a systematic comparison between the data-driven clusters identified through computational analysis and the theory-informed categories established during manual coding. While significant overlaps exist, the two methods provide unique insights that enhance the overall understanding of AI governance discourse.

The NVivo® analysis provided *contextual depth*, capturing explicit policy narratives, regulatory justifications, and ethical framings that are central to the European Commission's discourse on AI. Manually coded segments consistently emphasized the alignment of AI policy with European values, democratic governance, and ethical responsibility. Recurrent narratives of *trust*, *safety*, and *European Identity* illustrate how the Commission legitimizes technological innovation through moral and institutional principles.

The qualitative coding approach reveals explicit policy framings and discursive justifications, such as the EC's repeated assertion that AI policies align with European values, democratic governance, and ethical responsibility. For example, manually coded themes emphasize narratives of trust, safety, and European Identity, that are instead not explicitly discernible from the LDA topic clusters.

The topic modelling method detects word associations that may not be immediately apparent in manual coding. For instance, Topic 6 in LDA links "business", "service", "biotechnology", "record", "citizen", and "solution", suggesting a strong connection between AI adoption, the private sector, and

public-facing services. Similarly, Topic 7 introduces "robotic", "billion", "private", and "investment", pointing to an economic dimension of AI.

The findings suggest that the EC's discourse on AI revolves around:

1. *Healthcare AI applications and data governance* (confirmed by both NVivo and LDA Topic 1),
2. *Regulatory risk frameworks and liability* concerns (confirmed by NVivo and LDA Topic 9),
3. *Europe's strategic leadership in AI governance* (confirmed by NVivo and LDA Topic 4),
4. *AI's intersection with business and investment* (captured more explicitly by LDA Topics 6 and 7).

European Health Data Space (EHDS): Obstacles vs Benefits

This subsection provides the results of the computational text analysis – i.e. word frequency and topic modelling – concerning the dominant lexical patterns in the European Commission's communications on AI in healthcare. The results provide a basis for understanding the framing tendencies examined qualitatively in the following sections.

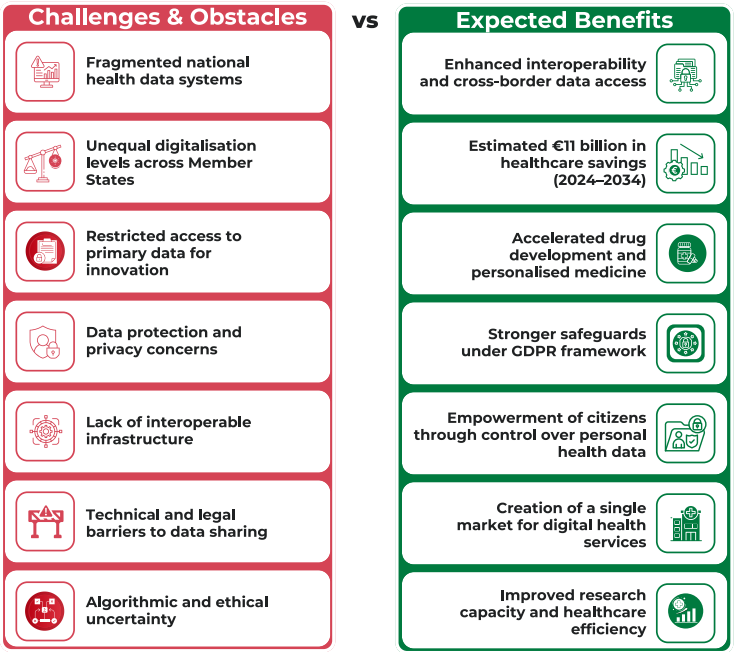
The comparison diagram between the "*Challenges & Obstacles*" and the "*Expected Benefits*" (**Errore. L'origine riferimento non è stata trovata.**) elucidates a stark dichotomy in the European Commission's discourse regarding the European Health Data Space (EHDS) and broader digital healthcare initiatives. The analysis provides a concise visual comparison of the obstacles and benefits most frequently identified in the corpus, thus illustrating the EC's balanced narrative. This narrative recognises systemic constraints yet frames digital health innovation as an ethical and social good. These findings are consistent with earlier analyses that identify data governance, trust, and ethical accountability as defining features of the EU's digital policy discourse (Ulnicane *et al*., 2021; Jasanoff, 2019). The prevalence of terms such as "health", "data", and "trust" underscores the Commission's positioning of AI within a socio-technical paradigm that prioritises public value. This distinguishes EU rhetoric from market-optimisation models observed in other contexts (Radu, 2021; Veale & Borgesius, 2021).

On the left side of the diagram, the coding references under "*Challenges & Obstacles*" indicate a range of structural and infrastructural issues that impede the full realisation of the potential of the EHDS. The recurrent themes of "*complex obstacles*" and the necessity for "*further digitalisation at the national level*" indicate that, despite a desire for cross-border health data exchange, many Member States still lack interoperable infrastructure. Moreover, the issue of restricted access to primary data for industrial use, coupled with the specific challenges associated with the collection, storage and utilisation of health data, reflect the ongoing difficulties encountered in the operationalisation of these systems.

In contrast, the "*Expected Benefits*" column presents the Commission's forward-looking narrative. The EHDS is expected to result in significant economic savings, with projections indicating that €11 billion could be saved over a ten-year period through enhanced access to health data and innovations in healthcare delivery and research (*Q&A on the European Health Data Space*, 24 April 2024). The benefits include accelerated drug development through enhanced data sharing, the advancement of personalised medicine, and the development of more sophisticated telemedicine capabilities. Additionally, there is a pronounced emphasis on the empowerment of individuals through the control of their health data, the fostering of a "*single market for digital health services*", and the improvement of research capabilities. The diagram summarises the prospective innovations in disease prevention, diagnosis, and treatment, along with the capacity to enhance healthcare efficiency and patient safety. It juxtaposes the aspirational outcomes of the EHDS against the real-world challenges that need to be addressed, thereby providing a clear illustration of the potential benefits and obstacles inherent in the implementation of this ambitious project. Although the projected benefits are considerable and have the potential to transform healthcare delivery in the EU, they are contingent on overcoming a number of entrenched obstacles related to infrastructure, privacy, and cross-border interoperability. The frequency and specificity of both the challenges and benefits coding indicate that the European Commission is actively engaged in addressing these issues while maintaining an optimistic outlook for the future of digital health in Europe.

This comparison serves as a visual representation of the EC's balanced narrative, which acknowledges the significant hurdles while also emphasising the transformative potential of digital health innovations under the EHDS framework.

Figure 2 – Challenges & Obstacles versus Expected Benefits of the European Health Data Space (EHDS)



Source: our elaboration based on coded data from European Commission communications (2018–2025)

General statements on AI: Issues and Benefits

The present subsection draws upon the quantitative patterns identified in Section 4.1 to explore the coalescence of qualitative codes into broader thematic categories, indicative of the manner in which the Commission articulates AI-related priorities and values in its communications.

The diagram comparing the benefits and issues surrounding AI in various speeches and documents (**Errore. L'origine riferimento non è stata trovata.**) provides qualitative elaboration and evolution of the quantitative results, visually synthesizing the benefits and issues associated with AI as expressed across Commission communications.. The following is a detailed commentary on the elaboration.

Issues: the left side of the diagram illustrates the pivotal issues and challenges inherent to the development of AI. These challenges are situated within the legal, ethical, and operational domains.

- *Ethical Concerns and Bias*: a considerable number of references indicate that AI presents a number of ethical challenges, including issues surrounding transparency, potential bias and discrimination. The potential for AI systems to replicate societal inequalities is a significant concern, as evidenced by the observation that women are less likely to receive job offers or that minorities are disproportionately targeted in predictive policing. These observations underscore the necessity for the formulation of ethical frameworks to mitigate the potential risks associated with biased outcomes of AI systems.
- *Liability and Legal Uncertainty*: the advent of AI has given rise to a number of crucial questions pertaining to liability, particularly in instances where AI systems have resulted in adverse outcomes. The lack of legal clarity for businesses, the disparate regulatory responses, and the opacity of AI algorithms are identified as significant obstacles to the adoption of AI. The intricate nature of AI systems renders it challenging to ascertain accountability, which can impede access to justice for those adversely affected by AI-driven harm, as evidenced by references to liability directives.
- *Trust and Transparency*: the lack of transparency in AI systems is a recurring theme, with speakers expressing concerns about the opacity of the decision-making processes employed by these systems. This phenomenon, which has been described as a “black box”, contributes to a broader deficit of trust among businesses and consumers. It is imperative that legal and regulatory measures are implemented to guarantee that AI systems operate in a transparent manner, thereby safeguarding fundamental rights and preventing unintended consequences.
- *Social and Economic Risks*: the potential impact of AI on employment is a significant concern, particularly the possibility of job losses due to automation. Small businesses often face

challenges in implementing AI, while larger economic actors must navigate intense competition in the global market. These socio-economic challenges are crucial and require support and resources to facilitate a smooth transition to AI-driven systems.

Benefits: the diagram's right side emphasises the numerous advantages of AI, particularly in domains such as healthcare, biomanufacturing, and broader societal applications. The principal themes are as follows.

- *AI in healthcare*: the presence of multiple coding references suggests that significant advancements in medical care are imminent, with AI set to improve diagnostics, personalised treatments and early disease detection. For instance, the capacity to analyse thousands of comparable medical images is highlighted as a notable advantage. One of the most notable advantages of AI is its capacity to assist medical professionals in making diagnoses in a more expeditious and precise manner.
- *Economic and Industrial Gains*: AI is regarded as a pivotal driver of innovation and industrial transformation. In their addresses, speakers have emphasised the capacity of AI to enhance operational efficiency in biomanufacturing, to inform financial decision-making, and to facilitate the development of new services in sectors such as agriculture and public administration. Furthermore, the economic impact of AI is evidenced by improvements in product and process optimisation across a range of sectors.
- *Environmental and Social Benefits*: AI is frequently cited as a means of addressing global challenges such as climate change, disaster forecasting, and cyberattacks. In addition to these, AI is also being deployed to manage health crises, optimise energy use, reduce traffic accidents, and enhance climate neutrality efforts. Furthermore, AI has the potential to assist marginalised communities, such as the elderly and disabled, and reduce inequalities in healthcare.
- *Personal Empowerment and Data Use*: the application of AI technologies has the potential to empower citizens in a number of ways. Firstly, it can facilitate greater control over one's own health data. Secondly, it can foster cooperation across different Member States. Thirdly, it can enable new opportunities in the field of personalised medicine. Another notable theme is the focus on data altruism, whereby citizens choose to contribute their data to support healthcare innovations.

The qualitative coding indicates that there is a framing of EU digital ethics discourse that mirrors previous studies respecting trustworthy AI acting as a regulative instrument and/or justification narrative (Floridi, 2019; Sartor, 2020). The constant link made to safety and an accountability of

democracy illustrates the EU's governing logic of justification for shared values (Schmidt, 2021), which signifies that it has made its AI submissive to its pre-existing governance structures.

European Commission's AI strategies alignment with broader European Identity (EI) politics

This section synthesises the previous findings by mapping dominant discursive patterns onto European identity politics. This interpretive synthesis links the textual evidence to the New Institutional Theory framework, demonstrating how AI policy discourse reflects the EU's broader normative identity.

The European Commission's AI strategies demonstrate a clear alignment with broader European Identity (EI) politics, emphasising adherence to core EU values, democratic principles and stringent data protection standards. The coding references consistently demonstrate how AI initiatives are integrated within the EU's ethical and legal frameworks, reflecting the Commission's commitment to the preservation of European values while fostering innovation.

1. The Commission reiterates on numerous occasions that AI development must adhere to the principles set forth by the EU (*Daily News 24 January 2024*, Press Release; *Commission launches AI innovation package to support Artificial Intelligence startups and SMEs*, Press Release, 24 January 2024). AI is presented as a means of bolstering European resilience and competitiveness while upholding the ethical standards that define the EU (*Tech and Geopolitics: Building European Resilience in the Digital Age* – Speech by Commissioner Thierry Breton, 5 September 2023). For instance, the *EU-US Trade and Technology Council Inaugural Joint Statement* (29 September 2021) emphasises the compatibility of AI with "*shared values and fundamental freedoms*", guaranteeing that the development and deployment of AI respect human rights and democratic governance. This reflects the EU's aspiration to spearhead global efforts in establishing a responsible AI ecosystem that echoes European values (*President von der Leyen's speech at the high-level opening session of the 2021 Digital Assembly: "Leading the Digital Decade"*, 1 June 2021).
2. The protection of data and the safeguarding of privacy appear to be of paramount importance. A significant aspect of the Commission's AI strategy is the focus on data protection and privacy. A considerable number of documents attest to this, including *Questions and Answers on the European Health Data Space* (24 April 2024) and *European Health Union: A European Health Data Space for People and Science* (3 May 2022). Moreover, the commitment to the protection of personal data ensures that AI initiatives are aligned with the EU's strict General Data Protection Regulation (GDPR), thereby embedding AI within a uniquely European approach to digital governance. This guarantees the maintenance of public

trust while facilitating technological advancement (*Daily News 24 January 2024*; *U.S.-EU Summit Joint Statement*, 20 October 2023).

3. The Commission's AI strategies position the EU as a global leader in establishing ethical and regulatory standards for AI, thereby reinforcing European Identity politics. For example, President von der Leyen's *2023 State of the Union Address* (13 September 2023) states that "*Europe has become the global pioneer of citizens' rights in the digital world*" illustrates how AI policies are positioned as extensions of European leadership in protecting citizen rights on a global scale. The references to the EU's partnerships with regions such as Africa and Latin America in advancing health products (*Commission takes action to boost biotechnology and biomanufacturing in the EU*, Press Release, 20 March 2024) demonstrate the Commission's aspiration to influence global AI governance through a framework of shared values and collaboration (*U.S.-EU Summit Joint Statement*, 20 October 2023). This approach aligns with the EU's broader global role as a proponent of ethical and responsible technology.
4. A recurrent theme is the balance between innovation, safety, and trust, which is central to the European Commission's approach to AI. The phrase "*in Europe, safety, trust, and innovation go hand in hand*" (*Artificial Intelligence: in Europe, Innovation and Safety Go Hand in Hand* – Statement by Commissioner Thierry Breton, 18 June 2023) encapsulates the Commission's vision of AI as a driver of economic and social progress that remains aligned with Europe's foundational principles of trust and safety. The equilibrium between technological advancement and ethical responsibility represents a fundamental aspect of the European Identity. AI initiatives such as the European Cancer Imaging Initiative (23 January 2023) are explicitly linked to the *European Strategy for Data* (*European Health Union: A European Health Data Space for People and Science*, Press Release, 3 May 2022).
5. The commitment to democratic values is a fundamental tenet of the Commission's approach. The Commission's strategies place an emphasis on the necessity for AI to operate within the bounds of "*shared democratic values*" (*U.S.–EU Summit Joint Statement*, 20 October 2023; *EU-US Trade and Technology Council Inaugural Joint Statement*, 29 September 2021). References to AI being developed in accordance with democratic principles reflect the broader alignment between AI strategies and the EU's commitment to strengthening democracy both within Europe and in its international partnerships (*President von der Leyen's speech at the 2021 Digital Assembly*, 1 June 2021).

A conceptual mapping was developed to further demonstrate how the EC's AI strategies align with the fundamental aspects of European Identity politics. The primary thematic categories found in the qualitative coding (such as ethics, data protection, and democratic governance) are connected to the

corresponding normative aspects of European identity, such as digital sovereignty, human rights, and participatory democracy, in this mapping, which is shown in *Appendix 1*.

These discursive alignments position the EC as an actor that not only regulates AI but also performs European Identity through its policy language. This finding lends further support to extant literature on the EU's "normative power" (Manners, 2002; Sjursen, 2006), demonstrating that the Union's approach to AI regulation simultaneously serves as an articulation of its unique political and ethical order. The findings indicate that AI policy communication functions as a mechanism for identity construction, in addition to its role in technological governance.

When considered as a whole, the analyses in Sections 4.3-4.5 demonstrate that the European Commission's discourse on AI incorporates regulatory, ethical, and identity-building dimensions. Beyond its technical content, this discourse performs the EU's broader political project: projecting a model of human-centric, value-driven technological governance. The study provides empirical evidence that contributes to the understanding of how institutional narratives maintain their legitimacy through the utilisation of moral vocabulary. The findings thus extend beyond mere description, thereby offering insight into the manner in which discourse itself constitutes a form of governance in the context of emerging technology policy.

Discussion

From the perspective of institutional theory, the recommendation guidelines on the governance of AI issued by the European Commission (EC) portray how the dimensions of legitimacy, compliance, and values intersect to create the digital landscape of European governance (RQ3). The risk-based approach advocated by the EC is predicated on the concept of *coercive isomorphism* (DiMaggio & Powell, 1983), wherein the categorisation of risk associated with the development of an AI system, particularly in high-risk sectors such as healthcare, serves as the foundation for the conversion of pressure from institutions into tangible instruments of conformity.

Simultaneously, *mimetic isomorphism* emerges through the adoption of regulatory sandboxes and testing facilities by the EC, mirroring policy instruments used in financial and data protection regulation. These sandbox mechanisms enabled controlled experimentation and knowledge exchange while maintaining public oversight, reflecting the EU's preference for experimentation within boundaries rather than deregulation. On the other hand, *normative isomorphism* is exemplified by the Commission's efforts towards ensuring the interoperability and cross-border data sharing, particularly through initiatives like the *European Health Data Space* (EHDS). This will promote the diffusion of the AI standards within the Member States and will solidify the collective commitment to transparency, accountability and trust.

The qualitative coding analysis confirms that indeed the EU's framing of the governance of AI is very much embedded within the mainstream discourse of digital ethics in the EU, in line with previous research on *Trustworthy AI* as a regulative instrument and a justificatory narrative (Floridi, 2019; Sartor, 2020). This rhetoric positions AI not as an autonomous sphere of innovation, but as a reality that may be tamed within and in accordance with prior European legal and moral frameworks. The constant emphasis on safety, fairness, and the accountability of democracy signals a logic of justification through shared values (Schmidt, 2006), in that AI must remain subordinate to the institutional norms that sustain European legitimacy.

In this regard, the EU's discourse aligns with the idea of Normative Power Europe (Manners, 2002), whereby governance is not only a matter of regulation but rather a performative act that projects the moral and ethical standards as a distinct global identity. Scholars have argued that this normative dimension is neither self-evident nor unproblematic (Sjursen, 2006), yet it remains crucial to the EU's endeavour to construct AI as a *public-value technology* rather than a purely market-driven innovation. Thus, the Commission's discourse embeds the development of AI in a sociotechnical imaginary (Jasanoff & Kim, 2015) that focuses on the progress of technology in the realm of democracy, citizens' trust, and egalitarianism – core features of the EU's identity politics in the digital era.

The *European Health Data Space* is one such example of this institutionalization where AI is integrated into a legal-ethical structure with inclusivity, data solidarity and public empowerment at its core. This initiative represents the Commission's value-driven approach, institutionalizing digital ethics as both governance logic and identity practice. This strengthens the EU identity as a normative entrepreneur who leads through ethics and morals instead of competing for liberalization.

However, a careful reading of the above analysis also discloses certain tensions. These stem from fragmentation between Member States, from a lack of ethical-legal enforcement, and from algorithmic biases. The reliance on a risk-based logic, while ensuring procedural accountability, risks reinforcing bureaucratic rigidity that may stifle innovation. In this sense, while the EC's AI policy discourse institutionalizes trustworthy AI as a legitimizing narrative, its effectiveness ultimately depends on whether these values translate into measurable social and regulatory outcomes.

In sum, the findings indicate that the EU's institutional strategy is both normatively ambitious and structurally constrained. Through coercive, mimetic, and normative mechanisms, the EC seeks to align AI innovation with European democratic ideals – simultaneously pursuing *governance legitimacy* and *identity construction*. This ambidextrous role of the EU highlights how technology regulation has become a values politics within this discursive practice of Europe, where AI regulation as a tool for reaffirming Europe's moral and institutional coherence.

Conclusions and Limitations

The present study demonstrates that European Commission communications on AI concurrently undertake governance and identity functions: they regulate technological innovation whilst constructing a narrative of European moral leadership. The analysis demonstrates that the Commission's discourse on AI is structured around four interrelated discourses, aligned with the EU's broader normative project: *ethical and legal regulation, innovation and competitiveness, data governance and trust,* and *European leadership and identity*.

While the research provides a comprehensive analysis of European Commission communications on AI, it is important to acknowledge the methodological limitations of the study. First, the corpus was limited to publicly available communications, which may not have captured internal deliberations or informal policymaking dynamics. Second, although the combination of computational and manual coding enhances validity, the hybrid approach potential methodological bias: topic modelling results depend on algorithmic parameterisation (e.g., number of topics and pre-processing steps), while qualitative coding remains subject to interpretive judgement despite intercoder checks. Third, the selection of English-language documents may underrepresent multilingual nuances in EU discourse or national-level rhetorical variations. Finally, the temporal scope of the study (2018-2025) reflects the evolution of AI policy discourse, but not its implementation outcomes. The methodological constraints thus delineate the boundaries of inference, while concomitantly suggesting directions for future comparative and longitudinal studies. Future work might extend the dataset to encompass multilingual and cross-institutional examples, employ a different set of computational methods (e.g., topic modelling via BERT), and examine cross-regional comparisons that can determine the applicability of a discursive framing of the EU. These findings underscore the need for continuous dialogue between policymakers, clinicians, and ethicists to ensure that AI governance frameworks remain both evidence-based and ethically grounded.

References

af Malmborg, F., Trondal, J. (2021). Discursive framing and organizational venues: mechanisms of artificial intelligence policy adoption. *International Review of Administrative Sciences*, 89(1), 39-58. https://doi.org/10.1177/00208523211007533

Almada, M., & Radu, A. (2024). The Brussels Side-Effect: How the AI Act Can Reduce the Global Reach of EU Policy. *German Law Journal*, *25*(4), 646–663. http://doi.org/10.1017/glj.2023.108

Barney, J. B. (1991). Firm Resources and Sustained Competitive Advantage. *Journal of Management*, 17(1), 99-120. https://doi.org/10.1177/014920639101700108

Barreto, I., & Baden-Fuller, C. (2006). To conform or to perform? Mimetic behaviour, legitimacy-based groups, and performance consequences. *Journal of Management Studies, 43*(7), 1559–1581. https://doi.org/10.1111/j.1467-6486.2006.00620.x

Battilana, J. (2006). Agency and Institutions: The Enabling Role of Individuals' Social Position. *Organization*, *13*(5), 653-676. https://doi.org/10.1177/1350508406067008

Battilana, J., Leca, B., & Boxenbaum, E. (2009). How actors change institutions: Towards a theory of institutional entrepreneurship. *The Academy of Management Annals*, 3(1), 65-107. https://doi.org/10.1080/19416520903053598

Birkstedt, T., Minkkinen, M., Tandon, A. and Mäntymäki, M. (2023). AI governance: themes, knowledge gaps and future agendas. *Internet Research*, 33(7), 133-167. https://doi.org/10.1108/INTR-01-2022-0042

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.

Boyd-Graber, J., Hu, Y., & Mimno, D. (2017). Applications of topic models. *Foundations and Trends in Information Retrieval*, 11(2-3), 143-296. http://dx.doi.org/10.1561/1500000030

Carriço, G. (2018). The EU and artificial intelligence: A human-centred perspective. *European View*, 17(1), 29-36. https://doi.org/10.1177/1781685818764821

Chan, H. S., Shan, H., Dahoun, T., Vogel, H., & Yuan, S. (2019). Advancing drug discovery via artificial intelligence. *Trends in Pharmacological Sciences, 40*(8), 592–604. https://doi.org/10.1016/j.tips.2019.06.004

Char, D. S., Abràmoff, M. D., & Feudtner, C. (2020). Identifying ethical considerations for machine learning healthcare applications. *American Journal of Bioethics, 20*(11), 7–17. https://doi.org/10.1080/15265161.2020.1819469

Coeckelbergh, M. (2020). *AI ethics*. Cambridge, MA: The MIT Press.

Crossnohere, N. L., Elsaid, M., Paskett, J., Bose-Brill, S., & Bridges, J. F. P. (2022). Guidelines for Artificial Intelligence in Medicine: Literature Review and Content Analysis of Frameworks. *Journal of medical Internet research*, *24*(8), e36823. https://doi.org/10.2196/36823

Dignum, V. (2019). *Responsible Artificial Intelligence. How to develop and use AI in a responsible way*. Cham: Springer.

Dimaggio, P. (1988). Interest and agency in institutional theory. In L. G. Zucker (Ed.), *Research on Institutional Patterns: Environment and Culture* Ballinger Publishing Co.

DiMaggio, P. J., & Powell, W. W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2), 147-160.

Dorado, S. (2005). Institutional Entrepreneurship, Partaking, and Convening. *Organization Studies*, *26*(3), 385-414. https://doi.org/10.1177/0170840605050873

European Commision. (2019). *Ethics guidelines for trustworthy AI*. Retrieved from https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

European Commission (2019). *Ethics Guidelines for Trustworthy AI*. Retrieved from https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai/

European Commission, 2018. Artificial Intelligence for Europe. COM 237. https://ec.europa.eu/transparency/regdoc/rep/1/2018/EN/COM-2018-237-F1-EN-MAINPART-1.PDF/

European Commission. (2018). *Artificial Intelligence for Europe.* European Union. Available at: https://digital-strategy.ec.europa.eu/en/policies/artificial-intelligence

European Commission. (2018). *Artificial Intelligence for Europe*. COM 237. Retrieved from https://ec.europa.eu/transparency/regdoc/rep/1/2018/EN/COM-2018-237-F1-EN-MAINPART-1.PDF/

European Commission. (2018). *Communication artificial intelligence for Europe.* Policy and legislation. Retrieved from https://digital-strategy.ec.europa.eu/en/library/communication-artificial-intelligence-europe

European Commission. (2019). Ethics guidelines for trustworthy AI. Publications Office of the European Union. Retrieved from: https://digital-strategy.ec.europa.eu/en/policies/regulatoryframework-ai

European Commission. (2021). 2030 digital compass: The European way for the digital decade [COM(2021) 118 final]. *Office for Official Publications of the European Communities*.

European Commission. (2022). *European Declaration on Digital Rights and Principles*. Retrieved from https://digital-strategy.ec.europa.eu/en/library/european-declaration-digital-rights-and-principles

Feldstein, S. (2024). Evaluating Europe's push to enact AI regulations: How will this influence global norms? *Democratization, 31*(5), 1049-1066. https://doi.org/10.1080/13510347.2023.2196068

Figueira, C. (2017). A Joint Communication to the European Parliament and the Council: towards an EU Strategy for International Cultural Relations, by the European Commission, 2016. *Cultural Trends*, *26*(1), 81–85. https://doi.org/10.1080/09548963.2017.1274371

Floridi, L. (2019). Establishing the Rules for Building Trustworthy AI. *Nature Machine Intelligence*, *1*, 1–2. https://doi.org/10.1038/s42256-019-0055-y

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., ... & Vayena, E. (2018). AI4People—An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines, 28*(4), 689-707. https://doi.org/10.1007/s11023-018-9482-5

Gerke, S., Minssen, T., & Cohen, I. G. (2020). Ethical and legal challenges of artificial intelligence-driven healthcare. In A. Bohr & K. Memarzadeh (Eds.), *Artificial intelligence in healthcare* (1st ed.). Elsevier. https://doi.org/10.2139/ssrn.3570129

Gray, G. C., Silbey, S. S. (2014). Governing Inside the Organization: Interpreting Regulation and Compliance. *American Journal of Sociology*, *120*(1), 96–145. https://doi.org/10.1086/677187

Greenwood, R., Oliver, C., Sahlin, K., & Suddaby, R. (Eds.). (2008). *The SAGE handbook of organizational institutionalism*. SAGE Publications.

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(suppl 1), 5228-5235.

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

Hardy, C., & Maguire, S. (2008). *Institutional entrepreneurship*. In R. Greenwood, C. Oliver, R. Suddaby, & K. Sahlin-Andersson (Eds.), The Sage Handbook of Organizational Institutionalism (pp. 198-217). SAGE Publications.

Hardy, C., & Maguire, S. (2017). Discourse, field-configuring events, and change in organizations and institutional fields: Narratives of DDT and the Stockholm Convention. *Academy of Management Journal, 53*(6), 1365–1392. https://doi.org/10.5465/amj.2010.57318384

Hyman D. A. (2001). Health care fraud and abuse: market change, social norms, and the trust "reposed on the workmen". *The Journal of legal studies*, *30*(2), 531–567. https://doi.org/10.1086/324674

Jacobi, C., van Atteveldt, W. H., & Welbers, K. (2016). Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, *4*(1), 89-106. https://doi.org/10.1080/21670811.2015.1093271

Jasanoff, S., & Kim, S. H. (Eds.). (2019). *Dreamscapes of modernity: Sociotechnical imaginaries and the fabrication of power*. University of Chicago Press.

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

Jurafsky, D., & Martin, J. H. (2021). *Speech and Language Processing* (3rd ed.). Pearson.

Krause, J. H. (2002). *Promises to keep: Health care providers and the civil False Claims Act*. *Cardozo Law Review, 23*, 1363. University of Houston Law Center No. 2007-A-19.

Krippendorff, K. (1980). Validity in content analysis. *Computerstrategien für die Kommunikationsanalyse*, *69*, 45p.

Lawrence, T. B., & Suddaby, R. (2006). Institutions and institutional work. *Handbook of organization studies*, 2, 215-254.

Leca, B., & Naccache, P. (2006). A Critical Realist Approach To Institutional Entrepreneurship. Organization, 13(5), 627-651. https://doi.org/10.1177/1350508406067007

Levi-Faur, D. (2011). Regulation and regulatory governance. Handbook on the Politics of Regulation, 3-22.

Liang, H. and Xue, Y. (2022). Save face or save life: physicians' dilemma in using clinical decision support systems. *Information Systems Research*, Vol. 33 No. 2, pp. 737-758. https://doi.org/10.1287/isre.2021.1082

Lysaght, T., Lim, H.Y., Xafis, V. and Ngiam, K.Y. (2019). AI-assisted decision-making in healthcare: the application of an ethics framework for big data in health and research. *Asian Bioethics Review*, Vol. 11 No. 3, pp. 299-314. http://doi.org/10.1007/s41649-019-00096-0

Maguire, S., & Hardy, C. (2006). The emergence of new global institutions: A discursive perspective. Organization Studies, 27(1), 7-29. https://doi.org/10.1177/0170840606061807

Manners, I. (2002). Normative Power Europe: A Contradiction in Terms? *Journal of Common Market Studies*, *40*, 235–258. https://doi.org/10.1111/1468-5965.00353

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to Information Retrieval*. Cambridge: Cambridge University Press.

Marcus, G., & Davis, E. (2019). *Rebooting AI: Building Artificial Intelligence we can trust*. New York: Pantheon Books.

Mayring, P. (2020). Qualitative content analysis: Demarcation, varieties, developments [30 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research, 20*(3), Art. 16. https://doi.org/10.17169/fqs-20.3.3343

Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340-363.

Milne, M. J., & Adler, R. W. (1999). Exploring the reliability of social and environmental disclosures content analysis. *Accounting, auditing & accountability journal*, *12*(2), 237-256.

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

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Science and engineering ethics*, *26*(4), 2141–2168. https://doi.org/10.1007/s11948-019-00165-5

Müggе, D. (2024). EU AI sovereignty: for whom, to what end, and to whose benefit?, *Journal of European Public Policy*, 31:8, 2200-2225, https://doi.org/10.1080/13501763.2024.2318475

Naik, N., Hameed, B.M.Z., Shetty, D.K., Swain, D., Shah, M., Paul, R., Aggarwal, K., Brahim, S., Patil, V., Smriti, K., Shetty, S., Rai, B.P., Chlosta, P. and Somani, B.K. (2022). Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility?. *Frontiers in Surgery*, Vol. 9, 862322. http://doi.org/10.3389/fsurg.2022.862322

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.

Patton, M. Q. (2014). *Qualitative research & evaluation methods: Integrating theory and practice*. Sage publications.

Polyviou, A., Zamani, E.D. (2023). Are we Nearly There Yet? A Desires & Realities Framework for Europe's AI Strategy. *Inf Syst Front*. 25, 143–159. https://doi.org/10.1007/s10796-022-10285-2

Powell, W. W., & DiMaggio, P. J. (1991). *The new institutionalism in organizational analysis*. University of Chicago Press.

Prakash, S., Balaji, J.N., Joshi, A. and Surapaneni, K.M. (2022). Ethical conundrums in the application of artificial intelligence (AI) in healthcare – a scoping review of reviews. *Journal of Personalized Medicine*, Vol. 12 No. 11, p. 1914. https://doi.org/10.3390/jpm12111914

Radu, R. (2021). Steering the governance of artificial intelligence: National strategies in perspective. *Policy and Society*, *40*, 1–16. https://doi.org/10.1080/14494035.2021.1929728

Saheb, T., Saheb, T., & Carpenter, D. O. (2021). Mapping research strands of ethics of artificial intelligence in healthcare: A bibliometric and content analysis. *Computers in biology and medicine*, *135*, 104660. https://doi.org/10.1016/j.compbiomed.2021.104660

Sartor, G. (2020). Artificial intelligence and human rights: Between law and ethics. *Maastricht Journal Of European And Comparative Law*, 27(6), 705-719. http://doi.org/10.1177/1023263X20981566

Schmidt, V. (2008). Discursive Institutionalism: The Explanatory Power of Ideas and Discourse. *Annual Review of Political Science*, *11*. https://doi.org/10.1146/annurev.polisci.11.060606.135342

Scott, W. R. (1995). *Institutions and organizations*. SAGE Publications.

Scott, W. R. (2008). Approaching adulthood: The maturing of institutional theory. *Theory and Society*, 37(5), 427-442.

Sjursen, H. (2006). The EU as a 'normative; Power: How can this be? *Journal of European Public Policy*, *13*. https://doi.org/10.1080/13501760500451667

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches. *Academy of Management Review*, 20(3), 571-610.

Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society*, 40(2), 137–157. https://doi.org/10.1080/14494035.2021.1928377

Tolbert, P. S., & Zucker, L. G. (1983). Institutional sources of change in the formal structure of organizations: The diffusion of civil service reform, 1880-1935. *Administrative Science Quarterly*, 28(1), 22-39.

Ulnicane, I., Knight, W., Leach, T., Stahl, B. C., & Wanjiku, W.-G. (2021). Framing governance for a contested emerging technology: Insights from AI policy. *Policy and Society, 40*(2), 158–177. https://doi.org/10.1080/14494035.2020.1855800

Veale, M., & Borgesius, F. (2021). *Demystifying the Draft EU Artificial Intelligence Act*. *22*. https://doi.org/10.48550/arXiv.2107.03721

Winfield, A. F., & Jirotka, M. (2018). Ethical governance is essential to building trust in robotics and artificial intelligence systems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 376*(2133), 20180085. https://doi.org/10.1098/rsta.2018.0085

Zhang, J., & Zhang, Z. (2023). *Ethics and governance of trustworthy medical artificial intelligence*. *BMC Medical Informatics and Decision Making, 23*, 103. https://doi.org/10.1186/s12911-023-02103-9

Zucker, L. G. (1986). Production of Trust: Institutional Sources of Economic Structure, 1840-1920. *Research in Organizational Behavior*, 8, 53-111.

Zucker, L. G. (1987). Institutional theories of organization. *Annual Review of Sociology*, 13(1), 443-464.

Chapter 3. **Technological Proximity and Patent Value in AI-Driven Healthcare: A Network-Based Innovation Analysis**

Simona Curiello[1, *], Enrica Iannuzzi[2], Claudio Nigro[2]

[1] Department of Social Sciences, University of Foggia, Italy

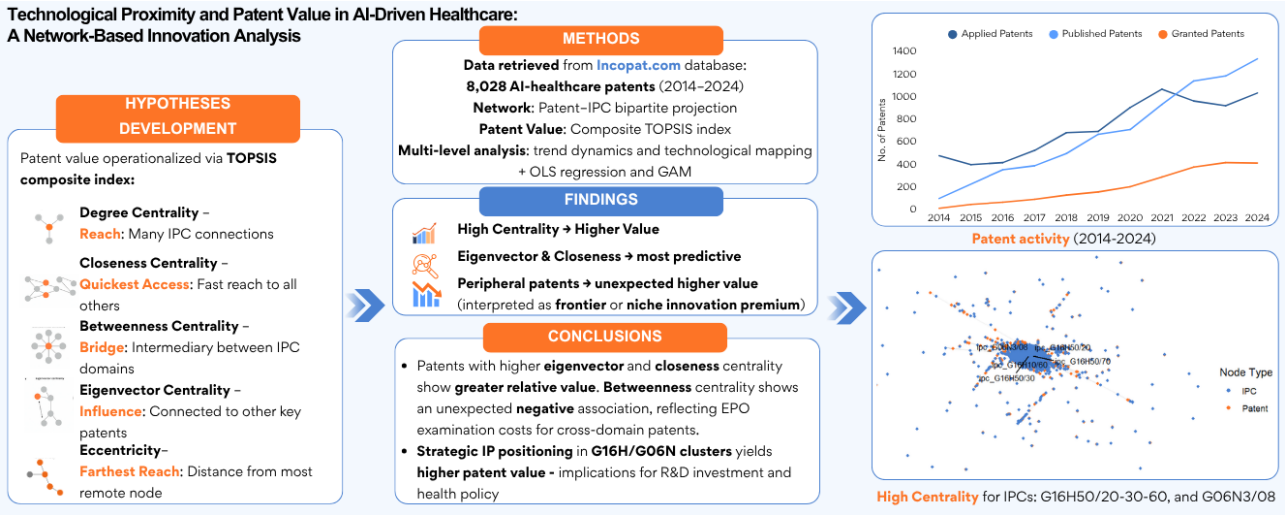
[2] *Department of Economics, University of Foggia, Italy*

[*] Corresponding Author. Email: simona.curiello@unifg.it

Abstract

This paper aims to explore how technological proximity within patent networks influences the strategic and economic value of Artificial Intelligence (AI) innovations in healthcare. Building upon open innovation and technological convergence theories, we conceptualize technological proximity as a mechanism that enables cross-domain knowledge recombination and impacts how innovation ecosystems create value. A bipartite network of International Patent Classification (IPC) co-occurrences has been constructed using 8,028 AI-healthcare patents filed between 2014 and 2024 from the IncoPat database, to capture technological proximity through multiple centrality measures. Adopting the TOPSIS methodology (Technique for Order Preference by Similarity to an Idel Solution) to combine multi-criteria value assessment, ordinary least square (OLS) regression models prove that patents occupying central and bridging positions within technology networks – particularly those with high betweenness and eigenvector centrality – exhibit greater economic and strategic value. Conversely, peripheral patents showing lower actual values may signal the rise of emerging technological domains. These results demonstrate that patent value is not solely a function of intrinsic technological quality of the innovation; rather, it also depends on structural embeddedness within interconnected innovation systems. This study contributes to innovation management research by theorizing how networked technological proximity mediates value creation in open innovation ecosystems, offering insights for policymakers, firms and R&D managers seeking to enhance innovation strategy in the rapidly evolving field of digital healthcare.

Keywords: Technological Proximity; Patent Value; Innovation Networks; Open Innovation; Artificial Intelligence; Healthcare; Network Centrality



Introduction

Artificial Intelligence (AI) is reshaping the global innovation landscape, transforming not only the boundaries of technology but also how organizations cooperate, coordinate, and compete. Evidence of this rapid evolution is provided by significant advancements in research and development (R&D) activities, as reflected in worldwide patent filings. For instance, the number of AI-related patent applications has increased dramatically over the past two decades. In the year 2000, approximately 10,000 AI-related patent applications were published, constituting around 6% of the total number of filings. By 2018, this total had more than doubled to over 60,000 annually, with the share of AI-containing patents growing from 9% to nearly 16% during the same period (Center for Security and Emerging Technology [CSET], 2020). This trend is also corroborated by the fact that the World Intellectual Property Organization's (WIPO) report that, between 1960 and early 2018, a total of nearly 340,000 AI patent families had been filed globally, and over half of them published since 2013 (Habibollahi Najaf Abadi & Pecht, 2020; Inside Global Tech, 2021). AI patent applications per annum had a growth rate of 6.5 between the years 2011 and 2017 alone, thereby indicating the acceleration of AI innovation (World Intellectual Property Organization [WIPO], 2019). Nowhere is this transformation more evident than in the healthcare sector, as AI-based systems that cover imaging, diagnostics, and clinical decision-making are redefining conventional value creation models (Benjamens *et al*., 2023; Topol, 2019).

Managerially, this steep escalation in patenting activity is indicative of more than just technological advancement; it reflects a strategic reorientation of innovation and competition. Innovation is shifting from isolated R&D initiatives towards networked value generation as businesses and research institutions increasingly engage in cross-domain recombination between the fields of biomedical sciences and information technology (Xin *et al*., 2021).

Patent analysis has become a standard tool in the field of AI for the measurement of technological trends, tracking the development of innovations, and studying the competition within corresponding domains (Fredström *et al*., 2021; Takahashi *et al*., 2024). Within this new paradigm, patents serve as strategic nodes in complex innovation ecosystems (Hwang *et al*., 2021). In addition to their legal function in safeguarding intellectual property, patents provide insight into the flow of knowledge across organizational and technological boundaries (Kim & Kim, 2018; Li *et al*., 2020). The capacity of firms to access complementary capabilities, form strategic alliances, and attain competitive advantage can all be influenced by their structural position within these technological networks. Such embeddedness is consistent with the open innovation logic (Chesbrough, 2003; Calvo *et al*., 2022), which claims that organizations generate more value when they open their boundaries to the inflow and outflow of ideas and take advantage of outside knowledge sources. However, despite the growing

significance of openness in technological development, the processes by which *technological proximity* – the relatedness between inventions – translates into patent value, are still conceptually poorly understood (Zhang & Ming, 2023). This conceptualization extends open innovation from a firm-level strategy to a network-structural mechanism, providing a novel theoretical bridge between micro-level knowledge recombination and macro-level innovation outcomes.

Drawing on theories of technological convergence (Bröring, 2010) and structural embeddedness (Granovetter, 1985), this study suggests that technological proximity within innovation networks represents the relational substrate through which open innovation unfolds (Broekhuizen *et al*., 2023). Because they integrate a variety of technological inputs and serve as connectors across previously unlinked domains, patents that hold more central or bridging positions within a technological proximity network are hypothesized to gain greater strategic and economic value (Teece, 1986). Conversely, peripheral patents, which are located further away from the technological core, might represent early or exploratory innovations that are vital for long-term transformation but frequently undervalued in the short term.

This study focuses on AI-driven healthcare as an ideal empirical setting to test these mechanisms. The selected domain exemplifies the convergence of multiple knowledge bases – information processing, medical instrumentation, and computational modelling – where innovation outcomes rely on cross-disciplinary cooperation and regulatory adaptation (Zhang & Liu, 2023). Unlike other, more mature technological fields, AI in healthcare operates under high uncertainty, heterogeneous knowledge bases, and intense regulation (Kannelønning *et al*., 2024; Comte *et al*., 2025; Broekhuizen *et al*., 2023). These characteristics heighten the importance of open innovation processes, whereby organizations must draw on distributed expertise, datasets, and computational methods across disciplinary and institutional boundaries (Allal-Chérif *et al*., 2023). The study thus offers unparalleled opportunities to examine how technological proximity and network embeddedness affect the creation and capture of innovation value in complex, interdependent systems.

The study employs a network-based methodology to evaluate the relationship between proximity and centrality metrics and patent value, operationalized through the TOPSIS method and regression models, by examining 8,028 AI-healthcare patents submitted between 2014 and 2024. Following the prior research published by Zhang & Ming (2023), this study aims to uncover patent trends in AI application in the healthcare system, track the evolution of AI technologies used in clinical practices and explore key players in AI innovation.

Accordingly, the paper addresses the following research question:

What effects does technological proximity have on the strategic and economic value of AI-healthcare innovations within patent networks? (Harhoff *et al*., 2003; Reitzig, 2003; Trappey *et al*., 2012)

By answering this question, the paper contributes to the literature in three important ways. First, it integrates open innovation and technological convergence theories into a single explanatory framework for understanding how knowledge proximity structures value creation in emerging innovation ecosystems. Second, it extends the network-based views of innovation by empirically linking patent centrality with market and strategic outcomes. Third, it provides managerial and policy implications for the R&D strategy, intellectual property management, and ecosystem governance in digital healthcare. Collectively, these contributions cut across conceptual and empirical insights to show that the value of innovation does not only lie in the invention per se but also in the network architecture through which technologies evolve, interact and generate competitive advantage.

Literature Review and Hypotheses

In the last decade, nations worldwide have strategically intensified investments in AI, including algorithms, AI-focused hardware, and sector-specific applications (World Intellectual Property Organization, 2019). Patents have been demonstrated to reflect an organisation's technological advancements and serve as indicators of R&D activities (Hufker & Alpert, 1994; Ernst, 2003). The information presented in patents offers valuable insights into the leading actors within a given technological field, providing a clear indication of their productivity (Pejic-Bach *et al*., 2019). Patent-network analysis has become a critical method for evaluating technological evolution across emerging domains, particularly in the field of AI. By examining patent data, researchers and policymakers can identify innovation trends, monitor the strategic direction of industries, and assess levels of R&D investment (Abbas *et al*., 2014; Fujii & Managi, 2018; Liu *et al*., 2021; Hain *et al*., 2022).

In the specific context of AI, patent analysis offers valuable insights into the pace of technological advancement, regional disparities in innovation output, and the actors – such as countries, corporations, and research institutions – driving development (Parteka & Kordalska, 2023).

Previous studies have utilized various classification systems to identify AI-related patents. For instance, Tseng and Ting examined AI-related patents from the USPTO dataset (Tseng & Ting, 2013), classifying them using the U.S. Patent Classification (UPC) system into four key subcategories: problem reasoning and solving, ML, network structures, and knowledge processing systems. Similarly, Fujii and Managi (2018) employed the International Patent Classification (IPC) system to categorize AI-related patents into four groups: biological models, knowledge-based models, specific

mathematical models, and other AI technologies. These classification methods provide a structured approach to analysing AI innovation and its evolution over time.

In an open innovation system, the value of patents plays a fundamental role in stimulating interest in R&D investment and innovation collaborations. High-value patents can, in turn, attract venture capital and push research to more enterprises, thus driving the development of related industrial clusters and making the healthcare sector smarter (Zhang & Ming, 2023).

In accordance with the principles of open innovation theory (Chesbrough, 2003), companies are increasingly turning to external knowledge sources for the purpose of enhancing their innovative performance. This change signifies that the generation of value is contingent not solely on the internal R&D capabilities of a company, but also on its capacity to access, absorb, and recombine external knowledge that constitutes a component of broader technological ecosystems (Moehrle & Gerken, 2012). From this standpoint, patents should not be regarded as independent entities, but rather as conduits through which knowledge is transferred between different fields of study (Cohen & Levinthal, 1990; Laursen & Salter, 2006; Hwang *et al*., 2021)). Open innovation may also be viewed as a structural phenomenon in network terms – firms or inventions that hold dominant positions within knowledge networks are able to leverage higher inflows and outflows of ideas, thus achieving superior innovation results (Powell *et al*., 1996; Guan & Liu, 2016).

While this study does not directly derive its hypotheses from open innovation theory, it makes an empirical contribution to it by demonstrating how network structures reflect the permeability of technological boundaries and knowledge flows – key constructs in open innovation paradigms (Chesbrough, 2003; Laursen & Salter, 2006; Powell *et al*., 1996). The present work deviates from the conventional approach of focusing on co-invention or social networks. Instead, it employs a novel methodology that utilises technological co-classification, an approach that has witnessed a surge in adoption within the domain of empirical patent research (Kim & Kim, 2018; Ko *et al*., 2021).

Building on structural network theory (Yan & Luo, 2017), we propose that a patent's position within the technological proximity network – modelled through IPC co-occurrence – has implications for its strategic and economic value.

Degree centrality, which is indicative of the number of direct links a patent has within the proximity network, has been shown to be associated with visibility and knowledge spillover. These phenomena, in turn, promote technology diffusion and commercialisation. Wanzenböck *et al*. (2014) emphasise that central nodes in European co-patent networks are often key innovation actors. Additionally, proximity to multiple knowledge domains has been demonstrated to increase the probability of forming valuable alliances (Wang *et al*., 2015; Hu *et al*., 2024). Consequently, the following hypothesis is proposed:

H1: *Degree centrality in the technological proximity network has a positive impact on patent value.*

Drawing from social network theory, the closeness centrality concept is traditionally applied to measure the relative closeness of actors in a network, or the extent to which a given node is embedded and interlinked (Zhao *et al*., 2024). In measuring technology networks, this measure specifies how much a patent is situated in close proximity relative to others and thereby provides a measure of its impact or influence potential.

Building on this, Aharonson and Schilling (2016) investigated how mechanisms like technological clustering, knowledge spillovers, and innovation patterns evolve with time through similar network structures. Analogously, Miranda and Borges (2019) demonstrated that closeness centrality determines the spread of information during initial stages of innovation, typically determining what ideas get noticed. While direct proof of higher patent value and proximity centrality is scarce, growing evidence proves its beneficial influence on more composite economic outcomes. While not as directly linked to immediate financial gain, studies suggest that this proximity measure supports faster technology dissemination and cross-sector relevance (Morescalchi *et al*., 2015; Miranda & Borges, 2019). Accordingly:

H2: *Closeness centrality has a positive impact on patent value due to enhanced information access and dissemination speed.*

Betweenness centrality is a metric used to quantify the control a patent exerts over the dissemination of information, thereby acting as a conduit between disparate domains. For intellectual property, generally it is not possible to have a single patent guard a complicated or technology-oriented good. As a result, companies tend to tap into a steady patent portfolio where individual patents reinforce each other to collectively cover the full technical range of a product (Zhang & Ming 2023; Li *et al*., 2025). When it comes to management of innovation, high betweenness centrality patents can play significant roles in strategic decision-making, particularly given that they are connecting points between otherwise disconnected technologies. This linking in turn contributes directly to patent value. As an example, a product can be based on two distinct technical aspects, each covered by patents of different types. Without a tie, the patents remain distinct in terms of legal or technical integration. But a focal patent that unites both elements can be a strategic hub, forming interdependencies and constructing a powerful patent barrier against competition.

Most of the theoretical justification of betweenness centrality is based on the theory of structural holes, which states that patents held in these central “bridge” positions are more valuable since they are at the centre of linking relationships across a network (Chang & Huang, 2016; Hu *et al*., 2023).

Technologically close, this implies that patents functioning as intermediaries within a portfolio are more influential and hence more valuable. Consequently:

H3: *Betweenness centrality in the proximity network positively influences patent value by positioning the patent as a bridge in the innovation landscape.*

Eigenvector centrality, which reflects a node's connection to other influential nodes, offers insight into a patent's indirect importance. Dong and Yang (2016) highlighted an essential distinction between eigenvector centrality and degree centrality, noting that while the former is most likely to advance new product development, the latter may be able to hinder it. In contrast, Kim (2019) noted that firms with high eigenvector centrality in industrial cluster networks will most likely experience stronger innovation outcomes. Similarly, Arranz *et al*. (2020) asserted that when firms are structurally embedded in networks through high eigenvector centrality, they can significantly enhance their performance on R&D projects. Linares *et al*. (2019) establish that patents connected to high-quality innovations often benefit from their neighbours' visibility and validation. This finding lends support to the following hypothesis:

H4: *Patents with high eigenvector centrality gain value through association with other high-impact technologies.*

Eccentricity in network analysis is a measure of the greatest distance from a particular node to any other node in the network. In many instances, the highest eccentricity of all the nodes determines the network's diameter. Nodes with high eccentricity values tend to be more isolated or peripheral in the overall structure.

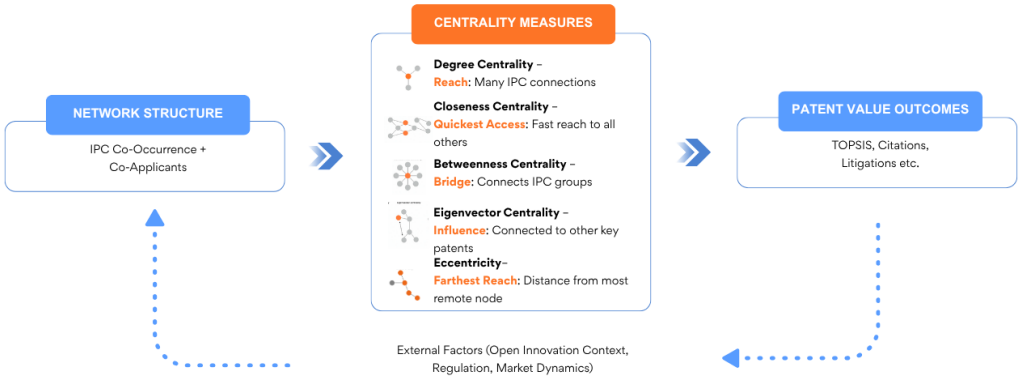
Within technological networks, Dotsika and Watkins (2017) showed how eccentricity analysis in keyword networks can facilitate the identification of emerging technological trends, especially those breaking away from mainstream research streams. In patent networks, eccentricity can be used as an indicator of the technological distance between a given patent and the mainstream, central pool of innovation. It is reasonable to hypothesise that emerging technologies may evolve into mainstream innovations; nevertheless, current patent valuations are predominantly informed by prevailing values, consequently resulting in assessments that frequently lag the rapid advancements of the technology sector. Hence:

H5: *Eccentricity has a negative effect on patent value in established domains, though may offer insights into emerging fields.*

The conceptual framework below shows how patent-technology proximity (through IPC co-occurrence and co-applicant structure) shapes several dimensions of network centrality, which in turn affect the relative patent value. To inform the structure and results of innovation networks, the model

incorporates external contextual factors, such as market demand, regulatory context, and open innovation dynamics, as moderators and sources of feedback.

Figure 1 – Conceptual Framework of centrality-value relationships in AI-healthcare patent networks



Source: our elaboration

Methodology

Drawing upon the prevailing existing literature and research design developed by Zhang and Ming (2023), this research examines the determinants of AI technology patent value. The investigation is informed by data sourced from the IncoPat database (https://www.incopat.com/).

To align this research more closely with to the organisational and clinical utilisation of AI technologies, a narrower search strategy was used, in contrast to the more general strategy employed by Zhang and Ming (2023). While their question cast a wide search net over a range of general categories of AI methods (e.g., "machine learning", "cloud computing", "semantic reasoning") and filtered over a wide range of health-related settings and IPC codes, our search targeted explicitly the use of AI in *clinical decision support systems* (CDSSs). In particular, this revised strategy prioritized terms directly denoting intelligent clinical devices – e.g., "computer-aided diagnosis", "rule-based systems", and "electronic health records" – to more accurately capture the practical and managerial integration of AI into healthcare workflow. The narrower timeframe (2014-2024) was also selected to best capture the latest decade of accelerated AI-healthcare convergence and to increase comparability with contemporary regulatory and commercial trends. This strategy provides a more domain-focused dataset (8,028 patents) to allow for deeper insight into the most prominent innovation patterns to clinical practice, organisational impact, and managerial approach:

((TIABC=("artificial intelligen*" OR "natural language process*" OR "information extrac*" OR "machine translat*" OR "natural language generat*" OR "semantic" OR "machine learning" OR "image* recogn*" OR "computer vision" OR "scene understand*" OR "speech recogn*" OR "cloud comput*" OR "pattern recogn*" OR "character recogn*" OR "video recogn*" OR "supervised learn*" OR "unsupervised learn*" OR "reinforced learn*" OR "multi-task learn*" OR "classification tree*" OR "regression tree*" OR "support

vector machine*" OR "artificial neural network*" OR "deep learn*" OR "logical learn*" OR "relational learn*" OR "probabilistic graphical model*" OR "rule learn*" OR "instancebased learni*" OR "latent representation" OR "bio-inspired approach*" OR "biometrics" OR "recommend* system" OR "logic program*" OR "fuzzy logic" OR "probabilistic reason*" OR "ontology engineer*" OR "search system" OR "expert system") AND (medic* OR drug OR iatric* OR treatment OR health OR pharmaceutic OR therapy OR diagnos* OR diseases OR sickness OR ailment OR illness OR pathem* OR prevalence OR epidemic OR infectious)) AND (IPC-LOW=(A61 OR G03 OR G16H OR G16B OR G16C OR G21F OR G03B15/14 OR B82Y5 OR G03B42/02)) AND (AD=[20050101 TO 20250409])) AND ((PNC=("EP")))

In this study, a two-mode undirected bipartite network was constructed to model the technological proximity relationships between patents and their associated IPC codes. The network was constructed utilising co-occurrence data from the "Patent ID" and "Main IPC" fields within the dataset.

Following a series of preprocessing steps (e.g., IPC splitting and prefixing node types), the resulting network topology comprised 10,389 nodes and 29,659 edges, where 7,746 nodes correspond to patents and 2,643 nodes correspond to IPC classifications. This network served as the foundational data structure for constructing the independent variables used in the analysis. Each edge in the network represents the linkage between a specific patent and an IPC classification, forming the analytical foundation for calculating patent centrality measures such as degree, closeness, betweenness, eigenvector centrality, and eccentricity.

To complement core centrality-based indicators, this study incorporates emerging metrics and control variables relevant to patent impact and strategic positioning. A co-applicant network was constructed by parsing joint applicant relationships from patent records, enabling the calculation of organizational level betweenness centrality for each patent based on its applicants' network position.

Additionally, control variables were included to account for contextual factors: 1) patent age, calculated as years since application, to control for time-based citation accumulation; 2) grant status, treated as a binary variable based on publication classification (granted vs. ungranted patents); 3) applicant type, distinguishing between firms, universities, and other assignees; and 4) applicant country, used as a proxy for regional legal or innovation system effects.

These variables were merged with centrality metrics and relative value scores derived from the TOPSIS model to build a multivariate regression framework assessing drivers of patent value.

Empirical Research Design

In the present study, the dependent variable (*relative patent value*) is operationalised based on structural attributes of patents within a proximity network. The topological position of a patent node reflects its significance and possible value. For operationalisation, a multi-criteria evaluation technique was applied, following the TOPSIS approach. This method employed several patent-level indicators, with the aim of generating a comparative value ranking for every patent within the given dataset. The ranking was then processed through a series of transformations involving reciprocals,

normalisation, and logarithmic scaling to ensure compatibility with ordinary least squares (OLS) regression. The specific measures utilized to estimate patent value included both bibliometric and commercial indicators, including citations' count, family size, and technology classes, combined with a composite value score retrieved from the IncoPat database.

The ultimate TOPSIS dataset included the following variables: simple and extended patent families' sizes, number of claims, word count of the first claim, IPC classification count, forward citations, litigation indicator (binary indicator), licensing activity, and assignment frequency. The goal of the current research was to determine how IncoPat scores are associated with patent retention for analysis. It was assumed that patents with non-zero IncoPat scores would be retained for analysis.

Table 1 – Dependent Variable Indicators

Variable Name (Weight)	Description	Indicator Usage
Number of simple family patents (0.056)	Count of patents that share the same priority	Reflects narrow protection (protection in fewer jurisdictions)
Number of expanded family patents (0.056)	Count of patents that are part of a broader family, i.e., share at least one priority	Reflects global coverage and broader filing strategy
Number of claims (0.056)	Total claims listed in the patent	Indicates technical breadth or legal strength
Words of first claim (0.056)	Word count of the first independent claim	May reflect complexity or specificity of main invention
Number of IPC (0.056)	Number of distinct IPC codes associated	Higher diversity → broader technological scope
Number of patents cited (0.056)	How many prior patents are cited by this one	Indicates knowledge dependency or novelty landscape
Number of lawsuits involved (0.056)	Number of legal disputes involving the patent	Indicates strategic or contentious value
Number of patent licenses (0.056)	Count of licensing agreements	Reflects commercial utility and market interest
Number of patent transfers (0.056)	Number of times the patent was assigned to a new owner	Reflects market dynamism, M&A, or patent trading value
System evaluation value (0.5)	A weighted composite score calculated by IncoPat	Aggregated indicator of economic and technological value based on above metrics and other proprietary ones

Source: our elaboration

Independent Variables

To investigate how the position of a patent within the technology proximity network affects its relative value, this research incorporated several centrality measures as independent variables into the regression analysis. The study employs a technology-class network, as opposed to a co-invention network, because its theoretical focus lies in knowledge recombination across technological domains, rather than interpersonal collaboration within inventor networks. In open innovation systems, value creation depends on the cognitive and technological relatedness of knowledge bases (Cantner & Graf, 2011). The network of International Patent Classification (IPC) co-occurrences has been demonstrated to be a valuable tool for analysing technological proximity – the extent to which different technological fields share complementary knowledge elements and evolve through recombination (Jaffe, 1986; Breschi *et al*., 2003). By operationalising technological proximity as the cognitive dimension of open innovation, the IPC network provides a direct and accessible approach to understanding the interconnections between technological developments. Conversely, co-invention networks have been demonstrated to be beneficial in terms of facilitating social proximity; however, they have been observed to be ineffective in capturing the relatedness between structural knowledge fields (Agrawal *et al*., 2008).

The centrality measures were designed to capture different dimensions of technological proximity and connectivity, as hypothesized in the conceptual framework. The selected measures, including degree, closeness, betweenness, and eigenvector centrality (Zhang & Ming, 2023), reflect both direct and indirect relational properties of a patent within the entire innovation network. Definitions and source citations for these variables are shown in Table 2.

Table 2 – Independent Variables

Analytical Dimension	Variable	Description	Key References
Network Centrality Characteristics	**Degree Centrality** (X1)	Measures the number of direct connections a patent node has with other nodes in the proximity network, indicating local connectivity	Sabidussi (1966); Freeman (1978); Hage & Harary (1995)
	Closeness Centrality (X2)	Reflects how close a patent node is to all other nodes by summing the shortest paths; a lower sum implies faster reachability across the network	Brandes (2001); Opsahl et al. (2010)
	Betweenness Centrality (X3)	Captures the extent to which a patent node acts as a bridge on the shortest paths between other nodes, indicating control over information flow	Freeman (1977); Borgatti (2005)
	Eigenvector Centrality (X4)	Assigns higher value to nodes that are connected to other influential nodes; captures recursive influence in the network	Bonacich (1987); Newman (2008)

Eccentricity (X5)	Measures the greatest distance between a patent node and any other node in the network; lower values indicate more central positioning	Wasserman & Faust (1994)

Source: our elaboration

Control Variables

To control for additional factors that might impact the relative value of patents, several control variables were included in the regression models. These variables' selection was informed by prior empirical research in the areas of innovation and patent research. For instance, patent's age, which is the number of years since the application date, is an important factor in predicting how much time passes before the patent has a chance to diffuse and accumulate citations. Prior studies have established that both patent citations and their relative significance increase over time (Hall *et al*., 2005; Lanjouw & Schankerman, 2004), thereby making the age of the patent a critical temporal factor to consider.

Grant status as a dichotomous variable involved attributing a value of 1 to granted patents (denoted by B1 or B2 codes). A patent grant is indicative of rigorous examination process completion and is commonly associated with high technological value as well as marketability (Gambardella *et al*., 2008; Mann & Underweiser, 2012).

Applicant type and country of origin were included to capture heterogeneity in innovation strategy and patenting behaviour. Finally, the count of International Patent Classification (IPC) codes, i.e., the number of IPC codes assigned to a patent, was added as an indicator of technological diversity or interdisciplinarity. Patents encompassing a greater number of IPC codes tend to cover more technical fields. This has been linked to higher recombinatorial potential and innovation impact (Lerner, 1994; Yan & Luo, 2017). The incorporation of these controls allows for the explanatory potential of the network centrality measures to be isolated by holding constant well-established determinants of patent value outlined in existing literature.

Findings

Trend Analysis

Focusing on the temporal dynamics of innovation in the domain of AI-driven healthcare, **Errore. L'origine riferimento non è stata trovata.** illustrates the distribution of patent applications, publications, and grants per year by the China National Intellectual Property Administration (CNIPA), the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), and the United States Patent and Trademark Office (USPTO) during the years 2014-2024. The condensed timeline reflects the period of acceleration in patenting in AI

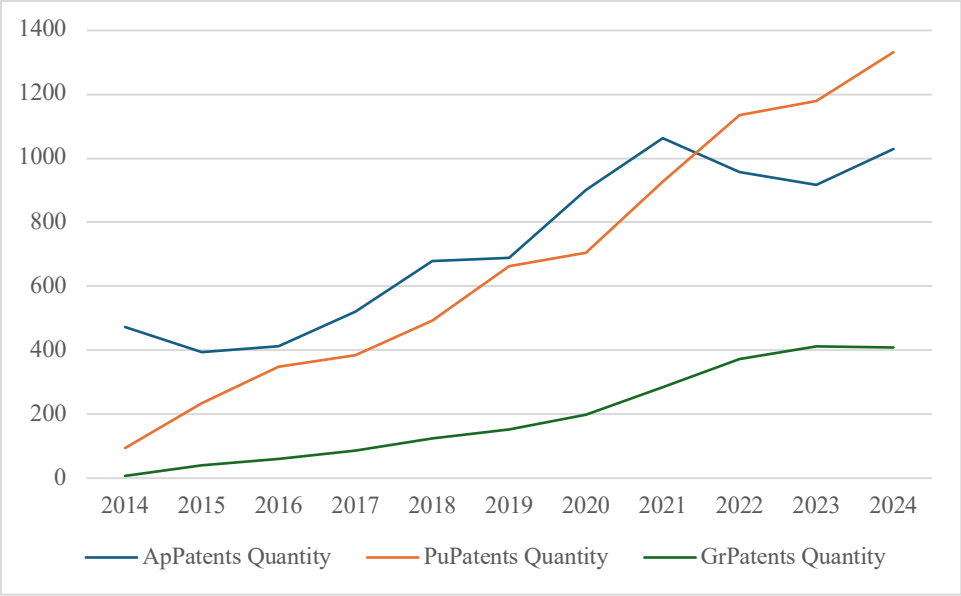
healthcare and describes the dynamics between the volume of applications, publication lag, and institutional authorisation.

The numbers show that patent applications (blue line) increased from 473 in 2014 to 1,063 in 2021, reflecting a compound annual growth rate (CAGR of 9.9%) over this period. This upward trajectory is in line with broader digital health trends and strategic investment in clinical AI. Although there was a slight dip after 2021, the application numbers remained robust at 1,029 applications in 2024, pointing to sustained industry activity despite market saturation or procedural delays.

Patent publications (orange line), which typically follow filings with a delay of 12–18 months, continued their strong growth, increasing from 94 in the year 2014 to 1,332 in 2024. This represents a robust CAGR of 28.6% and is a function of the maturation of filings during the 2018–2021 peak period. The consistent growth in published patents indicates that a larger share of earlier filings is entering the public domain and offering valuable information on disclosed innovation pathways.

The trend of granted patents (mapped through the grey line) delivers further strategic observations. From a modest number of 7 grants in the year 2014, this number steadily grew to an aggregate of 408 awarded patents in the year 2024. This signifies growth of over 5,700%, with a CAGR of 45.8%, reflecting a significant institutional shift in the validation and approval of AI-healthcare innovations. The year 2024, which recorded the highest ever grants, marks a tipping point for the regulatory and legal consolidation of previous innovations.

Figure 2 – Temporal Trends in Patent Activity



Source: own elaboration

Taken together, the data reveal a typical innovation lifecycle: a sharp influx of applications in the late 2010s, followed by a steady flow of publications and a significant growth in grant funding. The

patterns reflect a transition from exploratory development to legally protected intellectual property. While recent years have been characterised by some stabilization in new filings, the legacy effect of past innovations going through publication and approval stages indicates the maturing AI-healthcare ecosystem in Europe, fuelled by market demands and institutional learning.

Cluster Analysis

The text-mining-based clustering analysis revealed the dominant semantic fields of AI innovations in healthcare patents. At the first level, general themes such as “machine learning”, “trained models” and “medical imaging” emerged with regularity. The subsequent breakdown of these general themes at the second and third levels provided more specific applications such as “*diagnostic imaging*”, “*acquisition parameters*”, “*image processing*”, and “*cell-free analysis*”. This hierarchical organization of concepts supports the hypothesis that a significant portion of AI-healthcare patents involve image-based diagnostic support systems, particularly in clinical radiology and imaging workflows.

As shown in Table 2, clustering analysis by International Patent Classification (IPC) reveals a high concentration in healthcare informatics (G16H), electric digital data processing (G06F), AI models (G06N), medical diagnosis (A61B), and Information and Communication Technology (ICT) (GO6Q). The clusters are useful to underscore the central role of AI in the fields of patient data processing, management, and interpretation. This distribution suggests an increasing relevance of software-based diagnostic support and smart health systems in European patent applications.

Table 3 – Top IPC Classes in AI-Driven Healthcare Patents (2005–2025)

IPC Class	Patents Quantity
G16H – Healthcare Informatics Information and communication technology [ICT] specially adapted for the handling or processing of medical or healthcare data	2,999
G06F - Electric Digital Data Processing (computer systems based on specific computational models G06N) In this subclass, the following terms or expressions are used with the meaning indicated: "handling" includes processing or transporting of data; "data processing equipment" means an association of an electric digital data processor classifiable under group G06F7/00, with one or more arrangements classifiable under groups G06F1/00 to G06F5/00 and G06F9/00 to G06F13/00	2,134
G06N – Computing Arrangements Based on Specific Computational Model	1,79
A61B – Diagnosis; Surgery; Identification (analysing biological material G01N, e.g., G01N33/48). This subclass covers instruments, implements, and processes for diagnostic, surgical and person-identification purposes, including obstetrics, instruments for cutting corns, vaccination instruments, finger-printing, psycho-physical tests	1,068
GO6Q - Information and Communication Technology [ICT] Systems or Methods Specially Adapted for Administrative, Commercial, Financial, Managerial or Supervisory Purposes	918

Source: own elaboration

Applicant Technology Focus

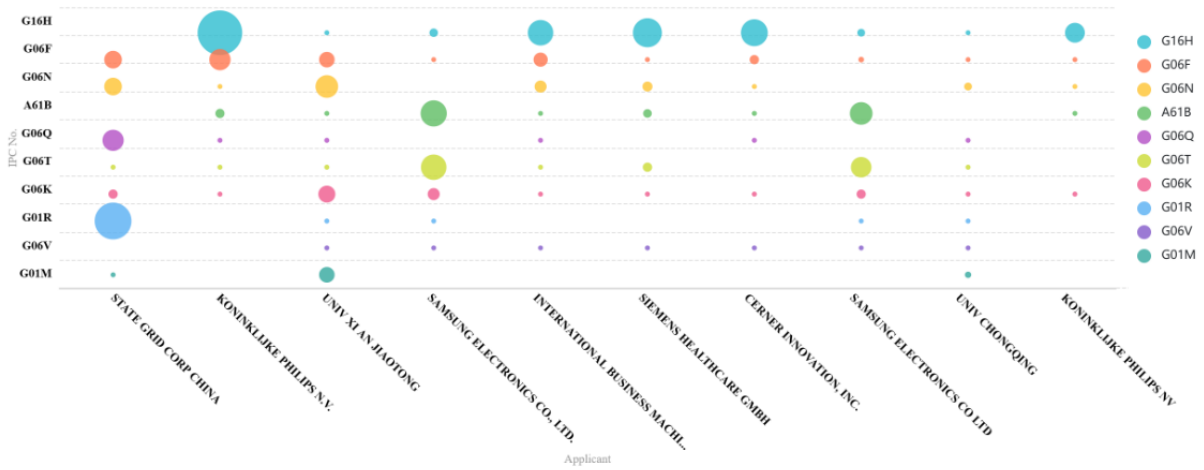
The figure below provides a detailed snapshot of applicant patent distributions across different IPC subclasses, detailing their technological focus, fields of specialisation, and innovation strategies. In particular:

- *Koninklijke Philips N.V.* has the most extensive and densest arrangement, with a particular leadership in G16H (Health Informatics) with 63 patents, and a balanced presence in G06F (Electrical Digital Data Processing), A61B (Diagnosis; Surgery), and G06T (Image Data Processing). This indicates a strategic push in AI-driven healthcare, imaging, and digital systems;
- *Samsung Electronics Co., Ltd.* is highly focused on A61B (37 and 32 patents in two listings), and also G06T, which points to intensive investment in medical imaging and diagnostic technologies;
- *International Business Machines Corp*. (IBM) demonstrates expertise in G16H, G06F, and G06N (Computer Systems Based on Biological Models), reflecting a high orientation toward AI algorithms and digital platforms for healthcare;
- *State Grid Corp China* is notable with its crushing dominance in G01R (Measuring Electric Variables) with 52 patents, reflective of its emphasis on energy measurement and grid-related technologies, but also shows consistent filings in G06F and G06N;

- Siemens Healthcare GMBH and Cerner Innovation Inc. show simultaneous investments in G16H, G06N, and A61B, reflective of their focus on clinical applications and informatics systems;
- Universities like *UNIV XI AN JIAOTONG* and *UNIV CHONGQING* are strongly represented in G06N and G06K (Recognition of Data; Presentation of Data), which indicates academic excellence in AI modeling and recognition technologies.

The subclass G06Q (Data Processing for Administrative or Commercial Purposes) is weakly represented among the applicants, indicating it is a less concentrated field in the AI-healthcare innovation landscape.

Figure 3 – Applicant Technology Focus



Source: own elaboration

Overall, this visualisation reveals disparities in the way that industrial participants and universities allocate R&D efforts across the IPC spectrum. Industrial stakeholders favour integration of health informatics and diagnostics, while universities specialise in fundamental AI algorithmic development and modelling.

Distribution of Technical Uses

The chart presents a hierarchical view of patent distributions across application domains (USE1Level), subdomains (USE2Level), and specific technologies or functions (USE3Level), revealing the depth and structure of innovation within the field of AI-healthcare. The "medical" domain leads all categories with 3,238 patents, with a significant focus on diagnostic methods (1,877) and diagnosis systems (541). Within these subcategories, notable concentrations of patents were observed, including 530 in the field of diagnosis, 248 in intelligent diagnosis methods, and 109–144

in diagnostic systems. These figures are indicative of substantial R&D activity in the realm of AI-driven clinical decision-making.

In a similar manner, the field of computational control (3,137 patents) is a central pillar, emphasising technologies such as computers (467), prediction (225), monitoring (268), and learning methods (239). These subdomains are further refined into advanced themes such as deep learning (105) and AI (103), signalling a foundational focus on data-driven models and predictive analytics in healthcare settings.

In the measuring experiments and process/method categories, key methods include detection (331 patents), analysis (326), and processing (282), which collectively demonstrate a pervasive emphasis on enhancing clinical accuracy and system intelligence. The mechanical equipment category, particularly support systems (226) and expert systems (109), suggests applications in smart devices and decision support tools.

Furthermore, the classification of diseases (1,463 patents), the communication of information (1,997) – including the processing of images and medical images (352, 98) – and electronic appliances (389) reflect the diversity of AI integration across both physical and digital healthcare tools.

The data demonstrate a concentration of patent activity in the domain of automated, predictive, and intelligent clinical support, with diagnostic tools and computational platforms representing the core of innovation.

Figure 4 – Patent Distribution across Domains

Source: own elaboration

Empirical Results

Table 4 presents the descriptive statistics for all key variables included in the analysis. Complete cases (n = 7,655) were used to calculate descriptive statistics for all relevant variables. To guarantee consistent comparisons across variables, observations with missing data were excluded from the analysis.

The dependent variable (Y), representing relative patent value, has a mean of 0.092 and a standard deviation of 0.029, indicating moderate variation across observations.

Among the centrality indicators, degree centrality (X_1) shows substantial variability with a mean of 779.63 and standard deviation of 746.61, reflecting large differences in node connectivity across the patent network. Closeness centrality (X_2) is more tightly distributed (mean = 0.457; SD = 0.080), while eigenvector centrality (X_4) has a mean of 0.135 and SD of 0.172, indicating moderate variation in influence-based positioning. Betweenness centrality (X_3) is notably low in both magnitude (mean = 0.00016) and variability (SD = 0.00036), which is common due to its sensitivity to rare bridging roles. Eccentricity (X5), reflecting distance from the most distant nodes, ranges from 1 to 9, with a mean of 5.85, capturing structural reach in the network.

Control variables include patent age (C_1), which varies between 1 and 11 years (mean = 5.02; SD = 2.98), and grant status (C_2), a binary variable with a mean of 0.069, indicating that approximately

7% of patents were granted. IPC diversity (C_3), measured as IPC_Count.x, has a mean of 3.85 with a standard deviation of 2.91, suggesting a moderate interdisciplinary spread.

Log-transformed versions of the network variables (log_ X_1 to log_ X_5) are also reported to reduce skewness and improve model interpretability. For example, log_degree has a mean of 5.76 and SD of 1.78, while log_betweenness remains close to zero (mean = 0.116; SD = 0.140) due to the original values' small scale. Other transformations such as log_eigenvector (mean = 0.00016) and log_eccentricity (mean = 1.91) similarly help stabilize variances in regression analysis.

Table 4 – Descriptive Statistics

Variable	N.	Mean	St. Dev.	Min.	Max.	Skewness	Kurtosis
Y	7655	0.09	0.03	0.01	0.46	0.42	7.9
X_1	7655	779.63	746.61	1	3411	0.82	-0.36
X_2	7655	0.46	0.08	0.17	1	2.07	16.48
X_3	7655	0	0	0	0.01	9.77	160.28
X_4	7655	0.13	0.17	0	1	1.38	1.26
X_5	7655	5.85	0.89	1	9	-1.89	9.48
log_X_1	7655	5.76	1.78	0.69	8.14	-1	0.13
log_X_2	7655	0.38	0.05	0.16	0.69	1.09	10.33
log_X_3	7655	0	0	0	0.01	9.73	159.21
log_X_4	7655	0.12	0.14	0	0.69	1.15	0.32
log_X_5	7655	1.91	0.17	0.69	2.3	-4.43	28.9
C_1	7655	5.02	2.98	1	11	0.44	-0.84
C_2	7655	0.07	0.25	0	1	3.39	9.51
C_3	7655	3.85	2.91	1	48	2.71	19.07

Source: our elaboration

Table 5 reports the correlation matrix of all variables used in this study with standardized Pearson correlation coefficients and significance levels. In this context, correlations of 0 to 0.5 represent low correlation, 0.5–0.8 reflect moderate correlation, and more than 0.8 represents high correlation.

As expected, degree centrality (X_1), closeness centrality (X_2), betweenness centrality (X_3), and eigenvector centrality (X_4) are all statistically significant and positively but weakly correlated with Y (r = 0.11*** to r = 0.17***), which indicates that higher centrality values tend to be associated with higher outcome values. In contrast, X_5 (eccentricity) has a tiny but highly significant negative correlation with Y (r = –0.06***), which indicates its theoretical distinction from measures of centrality based on distance.

Of the control variables, C_1 (time-based measure), C_2 (status-based), and C_3 (network complexity or number of categories) all exhibit statistically significant positive correlations with the dependent variable. Notably, the strongest correlation is for C_2 (r = 0.26*** with Y), followed by C_3 (r =

0.20***) and then C_1 (r = 0.12***), which means that these characteristics can have additive predictive capability in outcomes modelling.

There are highly significant inter-correlations among centrality measures, particularly between X_1 (degree) and X_4 (eigenvector; r = 0.97***), and between X_1 and X_3 (betweenness; r = 0.33***). Log-transformed versions reaffirm them, with log_X_1 and log_X_3 having high correlation (r = 0.72***) and between log_X_3 and log_X_4 (*r* = 0.25***), reinforcing concerns of multicollinearity. The correlation between X_2 (closeness centrality) and both X_1 (*r* = 0.53***) and X_4 (*r* = 0.46***) further suggests overlapping informational content among these measures.

To minimize multicollinearity and non-linearity, log transformations were applied to centrality variables, and model specifications were carefully crafted to prevent the simultaneous inclusion of highly correlated predictors, especially those with observed correlations greater than r = 0.70 (e.g., degree and eigenvector centrality). The shift from baseline models to log-transformed and non-linear specifications reflects these methodological decisions, which are further explained in the analytical strategy section.

Table 5 – Correlation Matrix

Var.	Y	X_1	X_2	X_3	X_4	X_5	C_1	C_2	C_3	log_X_1	log_X_2	log_X_3	log_X_4	log_X_5
Y	-													
X_1	0.17***	-												
X_2	0.11***	0.53***	-											
X_3	0.12***	0.33***	0.21***	-										
X_4	0.16***	0.97***	0.46***	0.25***	-									
X_5	-0.06***	-0.43***	-0.78***	-0.10***	-0.43***	-								
C_1	0.12***	-0.12***	-0.06***	-0.08***	-0.13***	0.03**	-							
C_2	0.26***	0.19***	0.09***	0.14***	0.18***	-0.07***	0.16***	-						
C_3	0.20***	0.45***	0.27***	0.59***	0.40***	-0.18***	-0.32***	0.16***	-					
log_X_1	0.15***	0.81***	0.51***	0.27***	0.70***	-0.28***	-0.11***	0.14***	0.39***	-				
log_X_2	0.12***	0.56***	1.00***	0.23***	0.49***	-0.77***	-0.07***	0.09***	0.29***	0.57***	-			
log_X_3	0.16***	0.98***	0.47***	0.25***	1.00***	-0.44***	-0.13***	0.19***	0.40***	0.72***	0.50***	-		
log_X_4	0.12***	0.33***	0.21***	1.00***	0.25***	-0.10***	-0.08***	0.14***	0.59***	0.27***	0.23***	0.25***	-	
log_X_5	-0.04***	-0.29***	-0.75***	-0.06***	-0.30***	0.96***	0.01	-0.04***	-0.11***	-0.10***	-0.73***	-0.30***	-0.06***	-

Source: our elaboration

Table 6 reports the results of a series of regression models investigating the relationship between network centrality metrics and relative patent value.

Models 1 through 5 are uncontrolled baseline specifications prevalent in conventional research practice across innovation network studies (e.g., Lee *et al*., 2015; Guan & Liu, 2016). This approach allows the direct observation of the main and interaction effects of centrality metrics without extraneous covariates.

Across the baseline models, degree centrality (X_1) consistently possesses a positive and statistically significant coefficient with respect to patent value (e.g., $\beta = 0.0000045$, $t = 2.27$, $p < .05$ in Model 1), corroborating Hypothesis 1. The eigenvector centrality measure (X_4) also possesses a strong negative effect across all specifications (e.g., $\beta = –0.206$, $t = –3.00$, $p < .01$ in Model 3), as anticipated by Hypothesis 4 and in agreement with earlier research on the traps of over-embedding (Ahuja, 2000). Betweenness centrality (X_3) is also significantly positive in certain models (e.g., β =

5.38, t = 5.11, p < .001 in Model 3), giving support to Hypothesis 3 empirically. However, closeness centrality (X_2) has marginal or non-significant effects (e.g., β = –0.121, t = –1.78 in Model 1), offering little support for Hypothesis 2. Eccentricity (X_5) was excluded from early models due to multicollinearity problems, as diagnosed using VIF analysis.

Models 2 through 5 examine interaction effects. Model 3 examines a degree × eigenvector interaction ($X_1 \times X_4$) with a moderated relationship, but the interaction effect is insignificant. Moreover, closeness × eccentricity ($X_2 \times X_5$) (Model 4) and betweenness × eigenvector ($X_3 \times X_4$) (Model 5) interactions are insignificant, showing minimal evidence for moderation effects.

Model 6 has interaction terms combined with main structural control variables like application year (C1), grant status (C2), and IPC count (C3). These control variables contribute heavily to the model (e.g., C1: β = 0.00520, t = 9.95, p < .001), while eigenvector centrality is significant at a marginal level (β = 0.00482, t = 1.80).

Model 7 incorporates log-transformed indicators of centrality to control for non-linear scaling relationships. In this model, log_eigenvector (centrality is significant (β = 0.00925, t = 2.57, p < .05), while others (e.g., log_degree and log_closeness) lose significance, possibly due to scale effects or interactions.

Model 8 incorporates the additional covariates and also a log_closeness × log_eccentricity interaction term in the model. Although the interaction fails to hold, log_eigenvector centrality (log_ X_4) remains significant (β = 0.00925, t = 2.57, p < .05), again confirming its persistent significance, as degree and closeness become insignificant — perhaps as a result of multicollinearity or conditional effects.

Model 9 adds all the predictive and control variables, confirming the persistent relevance of log_eigenvector (log_X_4) (β = 0.00925, t = 2.57) and non-relevance of log_degree (log_X_1) and log_closeness (log_ X_2), potentially pointing towards hidden multicollinearity or context-dependence.

Model 10 employs a *Generalized Additive Model* (GAM) with smoothed splines, offering a non-parametric formulation of the centrality–value relation. The significance of smooth terms (e.g., s(log_eigenvector): EDF = 6.74, p = 0.0098; and controls like C1: EDF = 31.121, p < .001 and C3: EDF = 14.96, p < .001) suggests nonlinear associations, highlighting the limitations of linear regression to detect complex network effects. This is in line with recent methodological advances in innovation studies, wherein GAMs and GAMMs are used to model subtle, context-specific dynamics (Wood, 2017).

Table 6 – Multiple OLS regression results

Var.	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
	OLS – Raw Centrality Terms	**OLS – Degree × Betweenness**	**OLS – Degree × Eigenvector**	**OLS – Closeness × Eccentricity**	**OLS – Betweenness × Eigenvector**	**OLS – Interaction + Controls**	**OLS – + X2, X5, and Interactions**	**OLS – Full Model (H1–H5 + controls)**	**OLS – Extended Controls**	**GAM – Smoothed Splines**
X_1	0.0000045*	0.0000047*	0.0000035	0.0000001	0.0000049*	0.0303	-0.1502			
	(2.27)	(2.37)	(1.58)	(0.06)	(2.44)	(0.09)	(-1.40)			
X_2	-0.121	-0.121	-0.102	-0.061	-0.121					
	(-1.78)	(-1.77)	(-1.47)	(-0.52)	(-1.77)					
X_3	5.06***	2.23	5.38***	5.29***	3.77*	-0.494	-0.582			
	(5.08)	(1.08)	(5.11)	(5.29)	(2.42)	(-1.42)	(-1.62)			
X_4	-0.106*	-0.106*	-0.206**	-0.114**	-0.105*	0.00482.	0.0112*	0.00925*		
	(-2.51)	(-2.42)	(-3.00)	(-2.58)	(-2.21)	(1.80)	(2.03)	(2.57)		
X5										
log_ X_1							0.00925*	0.00925*	0.00925*	(non-linear)
							(2.57)	(2.57)	(2.57)	1.862
log_ X_2						-0.493	-0.1502			
						(-1.47)	(-1.40)			
log_ X_3						0.079	0.014	-4.36	-4.36	(non-linear)
						(0.70)	(0.50)	(-1.09)	(-1.09)	14.96***
log_ X_4							-0.582		-0.582	(non-linear)
							(-1.62)		(-1.62)	0.0098**
log_ X_5							0.014		0.014	
							(0.50)		(0.50)	
C_1						0.00520***				(non-linear)
						(9.95)				31.121***
C_2						0.0503***				0.0441***
						(8.80)				
C_3						0.0032**				(non-linear)
						(3.14)				14.96***

Source: our elaboration

Discussions and Conclusions

In this study, an investigation on how technological proximity within patent networks influences the value of AI-healthcare innovations has been conducted. Over 8,000 patents data were retrieved from a robust dataset of worldwide patent filings over the past decade, and the findings illustrate a deep interdependency between the structural position of the patent in the technological landscape and its resultant economic and strategic potential. Drawing upon prior literature on the topic, centrality measures including degree, betweenness, and especially eigenvector centrality emerge as a consistent

indicator of patent value, revealing that patents do not operate in isolation, but there derives significance from their relationship and connectedness within broader innovation ecosystems.

The results validate the hypothesis that patents with more direct and influential connections – namely those that are conduits or find themselves in high-impact technological cluster – enjoy higher value. Interestingly, the fact that *eigenvector centrality* continues to be useful despite varying statistical models suggests that being connected to "influential neighbours" in the network has intrinsic value. This echoes broader theories of network embeddedness, where power is not only a question of direct affiliation, but also proximity to high-value knowledge spaces (Kim *et al*., 2019; Arranz *et al*., 2020).

Nevertheless, not all the network metrics translated clearly into patent value. Closeness centrality, for instance, had inconsistent effects, implying that while rapid access to information should theoretically be beneficial, it did not straightforwardly translate into economic value within the realm of patents. Similarly, eccentricity did not emerge to be a statistically significant predictor, though the theoretical significance in locating emerging or peripheral innovations is an area of study that would be worthwhile to explore further.

The results of the present study echo and expand upon prior research. In particular, the findings appear to be consistent with the work published by Zhang & Ming in 2023. The authors examined proximity effects in general industrial AI and found that degree and eigenvector centrality measures positively impact patent value, whereas betweenness and eccentricity may exert negative effects, especially under certain conditions. This difference reflects important alterations in domain-specific scenarios: in AI-healthcare, bridging roles (i.e., high betweenness) appear to maintain strategic value, possibly due to the interdisciplinary nature of clinical applications, where integration across knowledge domains is critical. Meanwhile, patents located on the periphery of the network – often dismissed in mainstream evaluations – may signal emerging innovations that are undervalued in the short term but pivotal in long-term transformation.

Furthermore, this study empirically contributes to open innovation theory by highlighting the importance of proximity-based mechanisms in patent value formation. High-centrality patents appear to be anchors in industrial clusters, facilitating knowledge diffusion and strengthening competitive advantage. As demonstrated by Zhang & Ming (2023), strategic IP layout can enhance innovation performance through both vertical integration and spillover effects. These structures support the growth of innovation ecosystems that are both resilient and capable of adapting to rapid technological change. Policy implications stream naturally from these insights. Policymakers are stimulated to support hierarchical IP strategies and incentivize the development of "core patents" that serve as *de facto* standards within emerging sectors. In addition, dynamic regulatory frameworks must be

developed to keep pace with technological change, particularly in complex domains like healthcare, where AI innovations often outpace ethical and legal readiness.

The results also point to the existence of a feedback mechanism: patents that gain value recognition through citations, litigation, or commercialization may subsequently draw more connections, increasing their future centrality, whereas high-centrality positions provide value through access to significant technological domains. This supports theories of cumulative advantage and preferential attachment in knowledge networks (Barabási & Albert, 1999) and is consistent with recursive dynamics seen in innovation systems.

Beyond statistics, our findings point to a management and policy issue: the AI-healthcare patent landscape is no longer an exceptional phenomenon. While China and the United States propel worldwide innovation and European actors carve out niches in informatics and diagnosis, patent terrain is being transformed into a battlefield as well as a frontier for cooperation. The growing engagement of universities, start-ups, and incumbent tech sector indicates a cross-dimensional relationship between business, academic, and institutional actors, each with their respective array of regulatory and ethical concerns regarding the deployment of AI within clinical practice.

Lastly, this study reaffirms the importance of network-driven R&D and IP management strategies. For innovation entrepreneurs, being in the right technological clusters and creating significant relationships between fields may be as essential as how new the technology itself is. For policymakers, this evidence indicates more responsive regulation mechanisms that acknowledge the accelerating nature and heterogeneity of AI-driven healthcare innovation.

Moving ahead, future research must extend the temporal and geographical horizons, add real-world rates of adoption, and examine the way network dynamics evolve over time. In doing so, we are then able to further uncover how the invisible architecture of innovation networks is shaping the very fabric of modern healthcare.

Limitations and Future Research

As with all empirical studies, the present one is subject to certain limitations. Firstly, while five key centrality measures were examined, the interaction effects among proximity dimensions – a focus in Zhang & Ming's (2023) work – remain underexplored in models under consideration. Secondly, while the study encompasses a comprehensive patent landscape, it is limited to patent data and therefore excludes market-level or clinical adoption metrics, which may influence or mediate patent value in real-world settings.

Future research could expand this work by integrating temporal dynamics, such as the evolution of centrality metrics over time and their influence on innovation lifecycle stages. Additionally, the

incorporation of alternative data sources – including licensing agreements, technology readiness levels, and citation velocity – could enrich the comprehension of patent valuation. Finally, future studies may benefit from advanced modelling approaches, including causal inference frameworks and agent-based simulations. These techniques have the potential to more accurately represent the intricate interdependencies characterising open innovation ecosystems in the domain of digital health.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used Deepl.com in order to improve the language. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

References

Abbas, A., Zhang, L., & Khan, S. U. (2014). A literature review on the state-of-the-art in patent analysis. *World Patent Information*, *37*, 3–13. https://doi.org/10.1016/j.wpi.2013.12.006

Agrawal, A., Kapur, D., & McHale, J. (2008). How do spatial and social proximity influence knowledge flows? Evidence from patent data. *Journal of Urban Economics*, *64*(2), 258–269. https://doi.org/10.1016/j.jue.2008.01.003

Aharonson, B. S., & Schilling, M. A. (2016). Mapping the technological landscape: Measuring technology distance, technological footprints, and technology evolution. *Research Policy*, *45*(1), 81–96. http://doi.org/10.1016/j.respol.2015.08.001

Ahuja, G. (2000). Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study. *Administrative Science Quarterly*, *45*, 425–455. https://doi.org/10.2307/2667105

Allal-Chérif, O., Costa Climent, J., & Ulrich Berenguer, K. J. (2023). Born to be sustainable: How to combine strategic disruption, open innovation, and process digitization to create a sustainable business. *Journal of Business Research*, *154*, 113379. https://doi.org/10.1016/j.jbusres.2022.113379

Arranz, N., Arroyabe, M. F., & de Arroyabe, J. C. F. (2020). Network embeddedness in exploration and exploitation of joint R&D projects: A structural approach. *British Journal of Management*, *31*(2), 421–437. http://doi.org/10.1111/1467-8551.12338

Barabasi, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science (New York, N.Y.)*, *286*(5439), 509–512. https://doi.org/10.1126/science.286.5439.509

Benjamens, S., Dhunnoo, P., Görög, M., & Mesko, B. (2023). Forecasting artificial intelligence trends in health care: Systematic international patent analysis. *JMIR AI, 2*, e47283. https://doi.org/10.2196/47283

Bonacich, P. (1987). Power and centrality: A family of measures. *American Journal of Sociology, 92*(5), 1170–1182. https://doi.org/10.1086/228631

Borgatti, S. P. (2005). Centrality and network flow. *Social Networks, 27*(1), 55–71. https://doi.org/10.1016/j.socnet.2004.11.008

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology, 25*(2), 163–177. https://doi.org/10.1080/0022250X.2001.9990249

Breschi, S., Lissoni, F., & Malerba, F. (2003). Knowledge-Relatedness in Firm Technological Diversification. *Research Policy*, *32*, 69–87. https://doi.org/10.1016/S0048-7333(02)00004-5

Breschi, S., Lissoni, F., & Miguelez, E. (2017). Foreign-origin inventors in the USA: testing for diaspora and brain gain effects. *Journal of Economic Geography*, *17*(5), 1009–1038. https://doi.org/10.1093/jeg/lbw044

Brin, S., & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. *Computer Networks and ISDN Systems, 30*(1–7), 107–117. https://doi.org/10.1016/S0169-7552(98)00110-X

Broekhuizen, T., Dekker, H., de Faria, P., Firk, S., Nguyen, D. K., & Sofka, W. (2023). AI for managing open innovation: Opportunities, challenges, and a research agenda. *Journal of Business Research*, *167*, 114196. https://doi.org/10.1016/j.jbusres.2023.114196

Buiten, M. C. (2019). Towards intelligent regulation of artificial intelligence. *European Journal of Risk Regulation, 10*(1), 41–59. https://doi.org/10.1017/err.2019.8

Burke, P. F., & Reitzig, M. (2007). Measuring patent assessment quality—Analyzing the degree and kind of (in)consistency in patent offices' decision making. *Research Policy*, *36*(9), 1404–1430. https://doi.org/10.1016/j.respol.2007.06.003

Calvo, N., Fernández-López, S., Rodríguez-Gulías, M. J., & Rodeiro-Pazos, D. (2022). The effect of population size and technological collaboration on firms' innovation. *Technological Forecasting and Social Change*, *183*, 121905. https://doi.org/10.1016/j.techfore.2022.121905

Cantner, U., Hinzmann, S., & Wolf, T. (2017). *The Coevolution of Innovative Ties, Proximity, and Competencies: Toward a Dynamic Approach to Innovation Cooperation* (pp. 337–372). https://doi.org/10.1007/978-3-319-45023-0_16

Center for Security and Emerging Technology. (2020). *Patents and artificial intelligence: A primer*. Georgetown University. https://cset.georgetown.edu/publication/patents-and-artificial-intelligence/

Chang, H., & Huang, M. (2016). The effects of research resources on international collaboration in the astronomy community. *Journal of the Association for Information Science and Technology*, *67*(10), 2489–2510. http://doi.org/10.1002/asi.23592

Chen, P., Xie, H., Maslov, S., & Redner, S. (2007). Finding scientific gems with Google's PageRank algorithm. *Journal of Informetrics*, *1*(1), 8–15. https://doi.org/10.1016/j.joi.2006.06.001

Chesbrough, H. W. (2003). *Open innovation: The new imperative for creating and profiting from technology*. Harvard Business Press.

Comte, V., Bertolini, L., Hulsman, R., Consoli, S., Leoni, G. et al. (2025). AI-driven Innovation in Medical Imaging - Focus on Lung Cancer and Cardiovascular Diseases. Publications Office of the European Union, Luxembourg. https://data.europa.eu/doi/10.2760/6281523

Dong, J. Q., & Yang, C. (2016). Being central is a double-edged sword: knowledge network centrality and new product development in U.S. pharmaceutical industry. *Technological Forecasting & Social Change*, *113*(1B), 379–385. http://doi.org/10.1016/j.techfore.2016.07.011

Dotsika, F., & Watkins, A. (2017). Identifying potentially disruptive trends by means of keyword network analysis. *Technological Forecasting and Social Change*, *119*, 114–127. http://doi.org/10.1016/j.techfore.2017.03.020

Ernst, H. (2003). Patent information for strategic technology management. *World Patent Information*, *25*(3), 233–242. https://doi.org/10.1016/S0172-2190(03)00077-2

European Patent Office. (n.d.). *CPC Cooperative Patent Classification browser*. Espacenet. Retrieved April 9, 2025, from https://worldwide.espacenet.com/patent/cpc-browser

Five IP Offices. (2023). *Home*. https://www.fiveipoffices.org/home

Fredström, A., Wincent, J., Sjödin, D., Oghazi, P., & Parida, V. (2021). Tracking innovation diffusion: AI analysis of large-scale patent data towards an agenda for further research. *Technological Forecasting and Social Change*, *165*, 120524. https://doi.org/10.1016/j.techfore.2020.120524

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry, 40*(1), 35–41. https://doi.org/10.2307/3033543

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. Social Networks, 1(3), 215–239. https://doi.org/10.1016/0378-8733(78)90021-7

Frerich, K., Bukowski, M., Geisler, S., & Farkas, R. (2021). On the Potential of Taxonomic Graphs to Improve Applicability and Performance for the Classification of Biomedical Patents. *Applied Sciences*, *11*(2), 690. https://doi.org/10.3390/app11020690

Fujii, H., Managi, S. (2018). Trends and priority shifts in artificial intelligence technology invention: A global patent analysis. *Economic Analysis and Policy, 58*, 60–69. https://doi.org/10.1016/j.eap.2017.12.006

Gambardella, A., Harhoff, D., & Verspagen, B. (2008). The value of European patents. *European Management Review*, *5*(2), 69-84. https://doi.org/10.1057/emr.2008.10

Granovetter, M. (1985). Economic action and social structure: The problem of embeddedness. *American Journal of Sociology, 91*(3), 481–510. https://doi.org/10.1086/228311

Gu, W., Wang, J., Zhang, Y., Liang, S., Ai, Z., & Li, J. (2024). Evolution of Digital Health and Exploration of Patented Technologies (2017-2021): Bibliometric Analysis. *Interactive journal of medical research*, *13*, e48259. https://doi.org/10.2196/48259

Guan, J., & Liu, N. (2016). Exploitative and exploratory innovations in knowledge network and collaboration network: A patent analysis in the technological field of nano-energy. *Research Policy*, *45*(1), 97–112. https://doi.org/10.1016/j.respol.2015.08.002

Habibollahi Najaf Abadi, H., & Pecht, M. (2020). Artificial intelligence trends based on the patents granted by the United States Patent and Trademark Office. *IEEE Access, 8*, 81633–81643. https://doi.org/10.1109/ACCESS.2020.2988815

Hage, P., & Harary, F. (1995). *Eccentricity and centrality in networks*. Social Networks, 17(1), 57–63. https://doi.org/10.1016/0378-8733(94)00248-9

Hain, D. S., Jurowetzki, R., Buchmann, T., & Wolf, P. (2022). A text-embedding-based approach to measuring patent-to-patent technological similarity. *Technological Forecasting and Social Change*, *177*, 121559. https://doi.org/10.1016/j.techfore.2022.121559

Hall, B. H., Jaffe, A. B., & Trajtenberg, M. (2005). Market value and patent citations. *RAND Journal of Economics, 36*(1), 16–38.

Harhoff, D., Scherer, F. M., & Vopel, K. (2003). Citations, family size, opposition and the value of patent rights. *Research Policy*, *32*(PII S0048-7333(02)00124-58), 1343–1363. http://doi.org/10.1016/S0048-7333(02)00124-5

Hu, F., Qiu, L., Xiang, Y., Wei, S., & Sun, H. (2023). *Spatial network and driving factors of low-carbon patent applications in China from a public health perspective*. Frontiers in Public Health. https://doi.org/10.3389/fpubh.2023.1121860

Hu, F., Shi, X., Wei, S., Qiu, L., Hu, H., Zhou, H., Guo, B. (2024). Structural Evolution and Policy Orientation of China's Rare Earth Innovation Network: A Social Network Analysis Based on Collaborative Patents. Polish Journal of Environmental Studies, 33(2), 1767-1779. https://doi.org/10.15244/pjoes/174511

Hufker, T., Alpert, F. (1994). Patents: A Managerial Perspective. Journal of Product & Brand Management, *3*(4), 44-54. https://doi.org/10.1108/10610429410073138

Hwang, J., Kim, B., & Jeong, E. (2021). Patent value and survival of patents. *Journal of Open Innovation*, *7*(2), 119. http://doi.org/10.3390/joitmc7020119

Inside Global Tech. (2021, February 12). *AI update: USPTO releases report on growth of artificial intelligence applications*. Covington & Burling LLP. https://www.insideglobaltech.com/2021/02/12/ai-update-uspto-releases-report-on-growth-of-artificial-intelligence-applications/

Jaffe, A. (1986). Technological Opportunity and Spillovers of R&D: Evidence from Firms' Patents, Profits and Market Value. *American Economic Review*, *76*, 984–1001.

Kamateri, E., Salampasis, M., & Perez-Molina, E. (2024). Will AI solve the patent classification problem? *World Patent Information*, *78*, 102294. https://doi.org/10.1016/j.wpi.2024.102294

Kannelønning, M. S. (2024). Navigating uncertainties of introducing artificial intelligence (AI) in healthcare: The role of a Norwegian network of professionals. *Technology in Society*, *76*, 102432. https://doi.org/10.1016/j.techsoc.2023.102432

Kim, H. (2019). How a firm's position in a whole network affects innovation performance. *Technology Analysis & Strategic Management*, *31*(2), 155–168. http://doi.org/10.1080/09537325.2018.1490398

Kim, S., & Kim, E. (2018). How intellectual property management capability and network strategy affect open technological innovation in the Korean new information communications technology industry. *Sustainability*, *10*(8), 2600. http://doi.org/10.3390/su10082600

Ko, S., Kim, W., & Lee, K. (2021). Exploring the factors affecting technology transfer in government-funded research institutes: The Korean case. *Journal of Open Innovation: Technology, Market, and Complexity*, *7*(4), 228. http://doi.org/10.3390/joitmc7040228

Lanjouw, J. O., & Schankerman, M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators. *Economic Journal, 114*(495), 441–465.

Lerner, J. (1994). The importance of patent scope: An empirical analysis. *RAND Journal of Economics, 25*(2), 319–333.

Li, L., Xin, D., Yang, X., & Yu, X. (2025). How does patent portfolio structure affect patent value in different technological environments? Evidence from Chinese high-tech industries. *Technological Forecasting and Social Change*, *215*, 124127. https://doi.org/10.1016/j.techfore.2025.124127

Li, S., Zhang, X., Xu, H., Fang, S., Garces, E., & Daim, T. (2020). Measuring strategic technological strength :Patent Portfolio Model. *Technological Forecasting and Social Change*, *157*, 120119. https://doi.org/10.1016/j.techfore.2020.120119

Linares, I. M. P., De Paulo, A. F., & Porto, G. S. (2019). Patent-based network analysis to understand technological innovation pathways and trends. *Technology in Society*, *59*, 101134. https://doi.org/10.1016/j.techsoc.2019.04.010

Liu, N., Shapira, P., Yue, X., & Guan, J. (2021). Mapping technological innovation dynamics in artificial intelligence domains: Evidence from a global patent analysis. *PloS one*, *16*(12), e0262050. https://doi.org/10.1371/journal.pone.0262050

Lu, Q., & Chesbrough, H. (2022). Measuring open innovation practices through topic modelling: Revisiting their impact on firm financial performance. *Technovation*, *114*, 102434. https://doi.org/10.1016/j.technovation.2021.102434

Mann, R. J., & Underweiser, M. (2012). A new look at patent quality: Relating patent prosecution to validity. *Journal of Empirical Legal Studies, 9*(1), 1–32. https://doi.org/10.1111/j.1740-1461.2011.01245.x

Miranda, M. G., & Borges, R. (2019). Technology-based business incubators: An exploratory analysis of intra-organizational social networks. *Innovation & Management Review*, *16*(1), 36–54. http://doi.org/10.1108/INMR-04-2018-0017

Moehrle, M. G., & Gerken, J. M. (2012). Measuring textual patent similarity on the basis of combined concepts: Design decisions and their consequences. *Scientometrics*, *91*(3), 805–826. https://doi.org/10.1007/s11192-012-0682-0

Morescalchi, A., Pammolli, F., Penner, O., Petersen, A. M., & Riccaboni, M. (2015). The evolution of networks of innovators within and across borders: Evidence from patent data. *Research Policy*, 44(3), 651–668. https://doi.org/10.1016/j.respol.2014.10.015

Morrison, G., Riccaboni, M. & Pammolli, F. (2017). Disambiguation of patent inventors and assignees using high-resolution geolocation data. *Scientific Data*, 4, 170064. https://doi.org/10.1038/sdata.2017.64

Mühlroth, C., Grottke., M. (2022). Artificial Intelligence in Innovation: How to Spot Emerging Trends and Technologies. *IEEE Transactions on Engineering Management*, *69*(2), 493–510. https://doi.org/10.1109/TEM.2020.2989214

Newman, M. E. J. (2008). Mathematics of networks. In L. N. Katz (Ed.), *The New Palgrave Dictionary of Economics* (2nd ed.). Palgrave Macmillan.

OECD. (2015). *Enabling the next production revolution: The future of manufacturing and services*. OECD Publishing.

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks, 32*(3), 245–251. https://doi.org/10.1016/j.socnet.2010.03.006

Palaniappan, K., Lin, E. Y. T., & Vogel, S. (2024). Global regulatory frameworks for the use of artificial intelligence (AI) in the healthcare services sector. *Healthcare*, 12(5), 562. https://www.mdpi.com/2227-9032/12/5/562

Parteka, A., & Kordalska, A. (2023). Artificial intelligence and productivity: Global evidence from AI patent and bibliometric data. *Technovation, 125*, 102764. https://doi.org/10.1016/j.technovation.2023.102764

Pejić Bach, M., Krstić, Ž., Seljan, S., & Turulja, L. (2019). Text Mining for Big Data Analysis in Financial Sector: A Literature Review. *Sustainability*, *11*(5), 1277. https://doi.org/10.3390/su11051277

Reitzig, M. (2003). What determines patent value? Insights from the semiconductor industry. *Research Policy*, *32*(PII S0048-7333(01)00193-71), 13–26. http://doi.org/10.1016/S0048-7333(01)00193-7

Sabidussi, G. (1966). The centrality index of a graph. *Psychometrika, 31*(4), 581–603. https://doi.org/10.1007/BF02289527

Seidman, S. B. (1983). Network structure and minimum degree. *Social Networks, 5*(3), 269–287. https://doi.org/10.1016/0378-8733(83)90028-X

Sharma, K., Manchikanti, P. (2024). Analysis of AI-Related Patents in Healthcare. In: Artificial Intelligence in Drug Development. Frontiers of Artificial Intelligence, Ethics and Multidisciplinary Applications. Springer, Singapore. https://doi.org/10.1007/978-981-97-2954-8_2

Takahashi, C. K., Figueiredo, J. C. B. de, & Scornavacca, E. (2024). Investigating the diffusion of innovation: A comprehensive study of successive diffusion processes through analysis of search trends, patent records, and academic publications. Technological Forecasting and Social Change, 198, 122991. https://doi.org/10.1016/j.techfore.2023.122991

Teece, D. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, *15*(6), 285–305. https://doi.org/10.1142/9789812796929_0001

Terry, N. (2019). Of regulating healthcare AI and robots. *Yale Journal of Law and Technology*, 21, 133–180. https://heinonline.org/hol-cgi-bin/get_pdf.cgi?handle=hein.journals/yjolt21&section=16

Tian, X., & Wang, J. (2018). Research on Spatial Correlation in Regional Innovation Spillover in China Based on Patents. *Sustainability*, *10*(9), 3090. https://doi.org/10.3390/su10093090

Topol, E. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Trappey, A. J. C., Trappey, C. V., Wu, C., & Lin, C. (2012). A patent quality analysis for innovative technology and product development. *Advanced Engineering Informatics*, *26*(1), 26–34. http://doi.org/10.1016/j.aei.2011.06.005

Tseng, C. Y., Ting, P. H. (2013). Patent analysis for technology development of artificial intelligence: A country-level comparative study. *Innovation*, *15*(4), 463–475. https://doi.org/10.5172/impp.2013.15.4.463

Wang, B., & Hsieh, C. (2015). Measuring the value of patents with fuzzy multiple criteria decision making: Insight into the practices of the industrial technology research institute. *Technological Forecasting and Social Change*, *92*, 263–275. http://doi.org/10.1016/j.techfore.2014.09.015

Wanzenböck, I., Scherngell, T., & Brenner, T. (2014). Embeddedness of regions in European knowledge networks: a comparative analysis of inter-regional R&D collaborations, co-patents and co-publications. *The Annals of Regional Science*, 53(2), 337–368. https://doi.org/10.1007/s00168-013-0588-7

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications.* Cambridge University Press. https://doi.org/10.1017/CBO9780511815478

Woerter, M. (2012). Technology proximity between firms and universities and technology transfer. *Journal of Technology Transfer*, *37*(6), 828–866. http://doi.org/10.1007/s10961-011-9207-x

Wood, S. (2006). Generalized Additive Models: An Introduction With R. In *Generalized Additive Models: An Introduction with R, Second Edition* (Vol. 66). https://doi.org/10.1201/9781315370279

World Intellectual Property Organization. (2019). *WIPO technology trends 2019: Artificial intelligence*. https://www.wipo.int/edocs/pubdocs/en/wipo_pub_1055.pdf

Xin, Y., Man, W., & Yi, Z. (2021). The development trend of artificial intelligence in medical: A patentometric analysis. *Artificial Intelligence in the Life Sciences*, *1*, 100006. https://doi.org/10.1016/j.ailsci.2021.100006

Yan, B., & Luo, J. (2017). Measuring technological distance for patent mapping. *Journal of the Association for Information Science and Technology, 68*(2), 423–437.

Yang, X., & Yu, X. (2019). Identifying Patent Risks in Technological Competition: A Patent Analysis of Artificial Intelligence Industry. *2019 8th International Conference on Industrial Technology and Management (ICITM)*, 333–337. https://doi.org/10.1109/ICITM.2019.8710719

Zaini, W. M. F., Lai, D. T. C., & Lim, R. C. (2022). Identifying patent classification codes associated with specific search keywords using machine learning. *World Patent Information*, *71*, 102153. https://doi.org/10.1016/j.wpi.2022.102153

Zhang, B., & Liu, X. (2024). Technology Proximity Mechanism and Collaborative Innovation Orientation: How to Coordinate Multiple Subsidiaries' Innovation Strategies? *Journal of the Knowledge Economy*, *15*(1), 706–731. https://doi.org/10.1007/s13132-023-01100-7

Zhang, B., & Ming, C. (2023). Patent value promotion based on the technology proximity network: An empirical analysis of artificial intelligence for healthcare. *Science, Technology & Society*, 28(2), 213–234. https://doi.org/10.1177/09717218231160441

Zhang, B., & Wang, H. (2021). Network Proximity Evolution of Open Innovation Diffusion: A Case of Artificial Intelligence for Healthcare. *Journal of Open Innovation: Technology, Market, and Complexity*, *7*(4), 222. https://doi.org/10.3390/joitmc7040222

Zhao, S., Wang, J., Ji, J., & Vincent Ekow, A. (2024). Proximity or alienation? Can knowledge type influence the relationship between proximity and enterprise innovation performance? *Technological Forecasting and Social Change*, *202*, 123314. https://doi.org/10.1016/j.techfore.2024.123314

Appendix 1. Coding Taxonomy

Code	Theoretical Foundation	Reference
Descriptive Framework on AI	Mimetic Isomorphism based on policy modelling due to uncertainty and legitimacy seeking	DiMaggio & Powell, 1983; Barreto & Baden-Fuller, 2006; Chan *et al*., 2019
Benefits	Advantages of AI, such as cost reduction and improved efficiency	Broussard, 2018; Fry, 2018; O'Neill, 2016
Issues	Ethical and operational issues arising from AI adoption	Borras and Edler 2014; Floridi *et al*., 2018
Unethical Uses	Examples of AI misuse or unethical practices	Floridi *et al*., 2018; Char *et al*., 2020
EU AI Act	Regulatory Institutionalization based on risk assessment and compliance mechanisms – Coercive Regulation	Levi-Faur, 2011; Gray & Silbey, 2014; European Commission, 2021
Risk Categorization	Categorization of AI based on risk level	
High-risk use cases	Critical applications with potential for significant societal impact	
Limited-Risk	Systems with moderate risks, requiring some oversight	
Minimal risk	Low-risk systems posing minimal harm	
Specific Transparency risk	AI systems with risks stemming from lack of transparency	
Unacceptable risk	Applications banned due to unacceptable societal risks	
European Identity	Mimetic Isomorphism based on policies modelling due to uncertainty and legitimacy-seeking through core EU values shaping AI policies and governance	Mügge, 2024
General Statement	Sector-specific rules exemplify institutionalization within formalized legal structures	Meyer & Rowan, 1977; von der Leyen, 2020
Health Data	Role of health data in driving AI innovation	Crossnohere *et al*., 2022
Challenges&Obstacles	Barriers and challenges in leveraging health data for AI	
Expected Benefits	Benefits of AI-driven health data use for diagnosis and research	
Pandemic's Digitalization	Digital acceleration in healthcare due to the COVID-19 pandemic	European Commission, 2021
Impacts	Overall impact of AI on healthcare and policy domains	Feldstein, 2024
Impact on Health Systems	AI's transformation of healthcare system operations	Chan *et al*., 2019
Impact on Healthcare Services	Effects of AI integration into healthcare services	
Impact on Heath Professionals' work	AI's influence on clinicians' workflows and practices	Reddy *et al*., 2020
Impact on Policy Makers	AI's impact on shaping policy decisions and regulations	Powell & DiMaggio, 1991
Member States Digital Level	Differences in AI adoption across EU member states	European Commission, 2021
Products&Applications	Applications of AI in drug discovery, imaging, and administration	Floridi et al., 2018
Regulatory Landscape	Regulatory Institutionalization through frameworks governing AI use	European Commission, 2021
European Digital Programmes	EU programs supporting digital transformation and AI development	Feldstein, 2024
Future directions	Future pathways for AI governance and innovation	Floridi *et al*, 2018

Other countries' framework	Comparative approaches to AI governance in other regions	Floridi *et al*, 2018
Security&Privacy	Privacy and security considerations in AI systems	Lawrence & Suddaby, 2006
AI Ethics and Transparency	Regulatory Institutionalization following the necessity of ensuring explainability and accountability in AI	Lawrence & Suddaby, 2006

Illustrative Data Excerpts Supporting the Four Discourses

To increase interpretive transparency, selected quotations from EU communications are provided below. These exemplify the language patterns that informed each discourse.

Discourse	Illustrative Quotation (Source File)	Interpretive Note
Ethical and Legal Regulation	"Europe must ensure that every AI system developed here is trustworthy and aligned with our values" (*President von der Leyen's Speech at the High-Level Opening Session of the 2021 Digital Assembly: "Leading the Digital Decade"*, 1 June 2021)	Reinforces EU's framing of AI ethics as integral to regulation
	"The new AI Liability Directive sets clear rules for transparency and accountability" (*Questions & Answers: AI Liability Directive*, 28 September 2022)	Shows the Commission's focus on binding legal frameworks
Innovation and Competitiveness	"InvestAI will mobilise €200 billion to strengthen Europe's position as a global hub for responsible innovation" (*EU Launches InvestAI Initiative to Mobilise €200 Billion of Investment in Artificial Intelligence*, 11 February 2025)	Links AI to economic growth and strategic resilience
Data Governance and Trust	"The European Health Data Space ensures that citizens' data are used safely and ethically to advance medical research" (*European Health Union: A European Health Data Space for People and Science*, 3 May 2022)	Demonstrates commitment to data ethics and trust in digital governance
European Leadership and Identity	"Europe has become the global pioneer of citizens' rights in the digital world" (*2023 State of the Union Address by President von der Leyen*, 13 September 2023)	Highlights the EU's identity as a normative power and value leader
	"Our partnerships with Africa and Latin America will advance responsible AI rooted in shared values" (*Commission Takes Action to Boost Biotechnology and Biomanufacturing in the EU*, 20 March 2024)	Illustrates the EU's external projection of its identity politics

Note: Quotations were selected from coded segments representing high-frequency nodes within each discourse. They exemplify the linguistic and conceptual patterns identified in the qualitative analysis.

Conceptual Mapping of AI Governance Themes and European Identity Dimensions
This table presents the conceptual mapping between the thematic categories identified in the qualitative analysis and the corresponding dimensions of European Identity (EI) politics.

Thematic Category	Illustrative Evidence (Source Documents)	Aligned EI Dimension	Interpretive Meaning
Ethical and Legal Adherence	*Daily News 24/01/2024*; *Commission Launches AI Innovation Package to Support Artificial Intelligence Startups and SMEs* (24 Jan 2024); *Tech and Geopolitics: Building European Resilience in the Digital Age* – Keynote Speech by Commissioner Thierry Breton (5 Sep 2023); *EU-US Trade and Technology Council Inaugural Joint Statement* (29 Sep 2021); *President von der Leyen's Speech at the 2021 Digital Assembly* (1 Jun 2021)	Human rights & fundamental freedoms	AI framed as compliant with the EU's democratic and ethical core values
Data Protection & Privacy	*Questions and Answers on the European Health Data Space* (24 Apr 2024); *Daily News 24/01/2024*; *U.S.-EU Summit Joint Statement* (20 Oct 2023); *EU-US Joint Statement of the Trade and Technology Council* (5 Dec 2022); *European Health Union: A European Health Data Space for People and Science* (3 May 2022)	Rule of law & digital sovereignty	GDPR alignment as an expression of lawful, citizen-centric governance
Global Leadership & Cooperation	*Commission Takes Action to Boost Biotechnology and Biomanufacturing in the EU* (20 Mar 2024); *U.S.-EU Summit Joint Statement* (20 Oct 2023); *2023 State of the Union Address by President von der Leyen* (13 Sep 2023)	Normative Power Europe	EU as global standard-setter promoting ethical and value-driven technology
Innovation, Safety & Trust	*Tech and Geopolitics: Building European Resilience in the Digital Age* – Keynote Speech by Commissioner Thierry Breton (5 Sep 2023); *Europe's Beating Cancer Plan: Launch of the European Cancer Imaging Initiative* (23 Jan 2023); *European Health Union: A European Health Data Space for People and Science* (3 May 2022)	Solidarity & social justice	"Trustworthy innovation" balancing competitiveness, public welfare, and patient safety
Democratic Values & Governance	*U.S.-EU Summit Joint Statement* (20 Oct 2023); *Joint Statement EU-US Trade and Technology Council of 31 May 2023 in Luleå, Sweden* (31 May 2023); *EU-US Trade and Technology Council Inaugural Joint Statement* (29 Sep 2021)	Democratic governance & participatory legitimacy	AI as an instrument for reinforcing democratic institutions and transatlantic cooperation

Example of Code Aggregation into Overarching Discourses

This appendix provides illustrative examples of how initial qualitative codes were merged into broader categories and then synthesized into the four overarching discourses described in Section 3 and visualised in Figure 1. The examples are representative and not exhaustive.

Overarching Discourse	Illustrative Code Categories	Example Initial Codes (NVivo Nodes)	Interpretive Focus / Meaning
1. Ethical and Legal Regulation	Regulatory frameworks, Ethical principles, Accountability mechanisms	"Ethical AI," "AI Act," "risk management," "trustworthy AI," "transparency," "rules for AI"	Emphasises the EU's commitment to ethical and legal oversight of AI systems
2. Innovation and Competitiveness	Economic growth, Digital innovation, Industrial leadership	"AI innovation," "startups," "investment," "digital transformation," "competitiveness," "future technologies"	Frames AI as a driver of European competitiveness and strategic autonomy
3. Data Governance and Trust	Data protection, Privacy, Digital sovereignty	"GDPR," "data spaces," "data sharing," "privacy," "citizens' trust," "security"	Highlights protection of personal data and the ethical management of data resources
4. European Leadership and Identity	Global leadership, European values, Democracy and human rights	"European model," "shared values," "human rights," "democracy," "global standards," "normative power"	Portrays the EU as a global leader promoting a value-based, human-centric approach to AI

Chapter 4. **Balancing Value and Risk: Clinicians' Perceptions and Adoption of AI-Enabled Clinical Decision Support Systems**

Simona Curiello[1,*], Adriane Chapman[2], Jeremy Wyatt[3]

[1] *Department of Social Sciences, University of Foggia, Italy*

[2] *Digital Health and Biomedical Engineering, University of Southampton, UK*

[3] *School of Primary Care, Population Sciences and Medical Education, University of Southampton, UK*

[*] Corresponding Author. Email: simona.curiello@unifg.it

Abstract

The increasing adoption of Artificial Intelligence (AI) in healthcare, particularly within Clinical Decision Support Systems (CDSSs), is transforming clinical practice and decision-making. Although AI-CDSSs hold the potential to improve diagnostic accuracy, operational efficiency, and patient outcomes, their implementation also creates ethical, technical, and regulatory concerns, affecting healthcare professionals' willingness to adopt these systems. Building on a value-based perspective, the study integrates the Unified Theory of Acceptance and Use of Technology (UTAUT) framework as determinants of perceived benefits and a risk-based perception model as determinants of perceived risks to develop a unified model exploring clinicians' behavioural intention to adopt AI-enabled CDSSs. A self-administered cross-sectional survey was distributed to licensed healthcare professionals to examine how validated factors influence perceptions of risks and benefits. Responses were collected from 215 clinicians across Italy and the United Kingdom. Recruitment was undertaken using email invitations, attendance at academic conferences, and direct approaches within healthcare settings. The findings offer theoretical and practical insights into human factors influencing AI adoption in clinical practice, underscoring the importance of value alignment, professional accountability and institutional readiness, and highlighting the need to foster clinician trust in AI tools beyond the boundaries of technical performance.

Introduction

In recent years, remarkable progress has been witnessed in the realm of AI-based healthcare applications, a topic that has been extensively documented (Senbekov *et al*., 2020; Asan & Choudhury, 2021; Rogers *et al*., 2021). Amongst these advancements, CDSSs have emerged as AI-based systems designed to help healthcare professionals make timely, evidence-supported, and data-driven decisions during diagnostic, prognostic, and treatment procedures. Typically integrated into Electronic Health Records (EHRs), CDSSs assist clinicians with real-time recommendations, warning against potential adverse drug reactions, and predicting patient outcomes based on historical data and clinical guidelines (Sutton *et al*., 2020). Extensive evidence shows the ability of AI-based systems to enhance diagnostic accuracy – for instance, in radiology and pathology applications that outperform human readers in image recognition (Jiang *et al*. 2017; Collado-Mesa *et al*. 2018; Esteva *et al*., 2019) – to improve treatment planning and optimization, such as individualized drug dosage calculation and therapeutic decision support (Topol, 2019; Liang & Xue 2022), and to support clinical interventions, including early warning systems for sepsis or adverse events that reduce human error and improve patient safety (He *et al*., 2019; Elhaddad & Hamam, 2024; Sendak *et al*., 2020).

With the continuous evolution of AI technologies, CDSSs are rapidly becoming a cornerstone of AI-driven transformation in healthcare with huge potential to improve care quality and patient safety (Topol, 2019). As these systems become more embedded in clinical practice, their influence extends beyond the operational efficiency to the very core of clinical judgment and patient outcomes (Liang & Xue, 2022). AI-driven decision-making can profoundly impact patients' lives in concrete and irreversible ways, particularly in diagnosis, treatment planning, and risk stratification sectors (Elhaddad & Hamam, 2024).

Although healthcare technology innovations have so far granted healthcare providers considerable advantages in the form of improved clinical efficiency and decision-making (Esteva *et al*., 2019; Topol, 2019; Mudgal *et al*., 2022), they have simultaneously produced unanticipated challenges, such as disruptions of cognitive processes (He *et al*., 2019), social interactional shifts (Siala & Wang, 2022), and workflow inefficiencies that can potentially introduce additional risks (Camilleri, 2024).

Emerging evidence suggests that even highly accurate clinical algorithms are not value-neutral; they are influenced by embedded assumptions, training biases, and structural constraints, which may give rise to ethical and moral dilemmas for both clinicians and patients (Martin, 2019; Murphy *et al*., 2021; Naik *et al*., 2022). These include opaque decision-making processes and limited explainability (Nazar *et al*., 2021), as well as challenges related to fairness, accountability, and trust in AI-assisted clinically informed decisions (Asan *et al*., 2020).

While healthcare professionals largely concur that AI has the potential to enhance patient outcomes, apprehensions regarding its ethical, organisational, and professional implications persist, with a majority not feeling adequately equipped to handle them (Collado-Mesa *et al*., 2018; Laï, *et al*., 2020). Empirical studies indicate that adoption is more likely when clinicians perceive the benefits – such as improved diagnostic precision, cost savings, or patient outcomes – to outweigh the risks (Choudhury, 2022; Meinert *et al*., 2018). Taken together, these findings suggest that clinicians' adoption of AI-based tools depends not only on technological performance, but also on their value assessments and professional judgments (Hendrix *et al*., 2022; van der Meulen, 2019).

However, despite the rapid development and deployment of AI technologies in healthcare, limited research has examined how clinicians themselves weigh the perceived benefits and risks of AI adoption. Existing studies remain disproportionally focused on technical efficacy, operational gains, or cost reduction (Overgoor *et al*., 2019; Seele *et al*., 2021), while neglecting the professional, ethical and psychological factors that shape clinicians' decision-making (Martin, 2019; Liang & Xue, 2022).

Although technology adoption models like the Technology Acceptance Model (TAM, Davis, 1989), the Unified Theory of Acceptance and Use of Technology (UTAUT, Venkatesh *et al*., 2003), the Technology Readiness Index (TRI, Lam *et al*., 2008), and the General Attitudes towards Artificial

Intelligence Scale (GAAIS, Schepman & Rodway, 2022) provide applicable insights into determinants of technology acceptance, they frequently fail to fully grasp the complex perceptions associated with AI risks, especially considering the intricate ethical, legal, and clinical contexts intrinsic to healthcare (Gagnon *et al*., 2016; Lee *et al*., 2025).

Decisions to implement new technologies in the healthcare industry are rarely driven solely by performance expectations. Instead, they represent a value-based evaluation in which clinicians weigh perceived benefits (e.g. enhanced accuracy, efficiency, or patient outcomes) against potential risks (e.g. liability, loss of autonomy, or ethical issues) (Kim *et al*., 2007). Building on this value-based perspective, this study seeks to offer more insightful understandings of the facilitators or barriers to the adoption of AI-embedded CDSSs in clinical practice by addressing the following research questions through a multivariate model that integrates the UTAUT framework as a determinant of perceived benefits and the risk-based approach (Esmaeilzadeh's model) as antecedents of perceived risks:

RQ1. *How do healthcare professionals perceive the benefits of AI medical devices, and in what ways do these perceptions influence their intention to adopt CDSSs?*

RQ2. *What are the risks perceived by healthcare professionals about the adoption of AI medical devices, and how do these risks influence their intention to use such technologies?*

Having established the study's rationale and research aims, the following section reviews prior empirical studies on AI adoption among clinicians to frame the research within existing academic discourse.

Related Work (Empirical Overview)

The growing body of empirical research on AI in healthcare reveals a nuanced blend of enthusiasm and apprehension among clinicians. A substantial number of studies highlight enduring concerns regarding its ethical, professional and organisational implications (Romero-Brufau *et al*., 2020; Turja *et al*., 2020).

For instance, Pedro *et al*. (2023) and Choudhury and Shamszare (2023) suggest that despite optimism from medical specialists on AI, scepticism remains on the impact of AI on healthcare quality and clinical procedures. Such fears are heightened by ethical considerations deriving from AI deployment, as highlighted by Petersson *et al*. (2023), who identify the importance of medical ethics in AI adoption. Similarly, Choudury *et al*. (2022) put emphasis on the need for training and accountability in the implementation of AI, highlighting reduced professional autonomy and potential biases. Studies by Vo *et al*. (2023) and Gillan *et al*. (2019) identify additional perceived risks such as data privacy, algorithm bias, and professional role modification, emphasizing the need for AI

validation and transparency. Finally, Valente *et al*. (2022) and Thekdi *et al*. (2023) focus on the necessity for explainable AI and appropriate employment of risk principles to facilitate patient safety and professional trust in AI-based applications.

Several investigations have also examined how different professional and demographic groups perceive AI in healthcare. For instance, a study by Pinto Dos Santos *et al*. (2019) quantitatively assessed the attitudes of medical students towards AI, revealing a consensus that AI would significantly enhance the medical field. In the same vein, Gong and colleagues (2019) carried out a study among Canadian medical students, discovering that a large majority of them (67%) felt AI would reduce the need for traditional medical positions, with 47% worried about the future of physicians given AI progresses. A 2019 survey held by the European Society of Radiology indicated reluctance to endorse the adoption of AI-based reports alone, with over half of the participants voicing strong reservations. Broadbent *et al*. (2012) examined attitudes towards healthcare robots within the setting of a retirement community and found a generally positive reception by the residents for them, contrasting with the views of staff and family members. In their comparison of human and AI-assisted pneumonia diagnoses, Patel *et al*. (2019) showed that collaborative and AI-deep learning models achieved superior accuracy to human experts alone.

Overall, the literature reflects a dual narrative in clinicians' responses to AI: optimism regarding its potential benefits coexists with apprehension about professional risk, liability and ethical accountability (Shamszare & Choudhury, 2023). This tension underscores the need for theoretical integrative models to explain how healthcare professionals balance perceived advantages and threats when evaluating AI tools.

To better understand these contrasting attitudes, scholars have increasingly resorted to theoretical models of technology acceptance. Among these, the Unified Theory of Acceptance and Use of Technology (UTAUT) developed by Venkatesh *et al*. (2003) has proven particularly influential in explaining the behavioural intentions that drive the uptake of healthcare innovations. Whilst the preceding studies highlight clinicians' perceptions of AI's benefits and risks, the UTAUT framework provides a structured lens to examine the psychological and organisational factors that shape adoption decisions. The theory identifies four core determinants of technology adoption: 1) *performance expectancy*, the extent to which users believe that the technology will improve job performance; 2) *effort expectancy*, the perceived ease of use; 3) *social influence*, the degree to which individuals feel that important people in their lives believe they should use the technology; and 4) *facilitating conditions*, the belief that organisational and technical infrastructures support usage.

A substantial body of literature has employed UTAUT framework to examine acceptance of technological innovations in the healthcare industry, including EHRs and mobile health systems

(Garavand *et al*., 2019; Ifinedo, 2012; Kim *et al*., 2016; Rouidi *et al*., 2022; Wills *et al*., 2008). Holden and Karsh (2010) used the model to investigate EHRs adoption, identifying *performance expectancy* as a key determinant of use. In a similar vein, Gagnon *et al*. (2016) highlighted the influence of *effort expectancy* and *social influence* on the acceptance of health information technology. Overall, the adoption of the UTAUT framework enables healthcare leaders to articulate targeted strategies that promote the effective implementation of innovative medical technologies.

Beyond these applications, this theoretical framework has also been employed to explore a variety of new digital health technologies, such as telemedicine and AI-based diagnosis. For instance, Sykes *et al*. (2009) found that the availability of *facilitating conditions* was the most influential factor for the adoption of telemedicine, while Cimperman *et al*. (2016) confirmed the continuing relevance of *social influence* in encouraging e-health services use among older adults.

Furthermore, *training adequacy* consistently emerges as a critical determinant of successful health technologies uptake. Inadequate user training and limited infrastructural support have frequently been cited as barriers to adoption, particularly in nursing homes and long-term care settings (Ko *et al*., 2018). A comprehensive systematic review by Gagnon *et al*. (2012) also underlines that proper user training is one of the organisational determinants of technology adoption among medical professionals. Similarly, Alrahbi *et al*. (2021) highlight the fundamental role of trained experts in the successful implementation of healthcare IT, thereby underscoring the necessity for skill development initiatives.

However, as UTAUT principally concentrates on utilitarian determinants of technology use, it may not fully capture the ethical and risk-related dimensions that characterise AI in clinical contexts in a comprehensive manner. Recent research has highlighted that medical practitioners still have significant concerns about the use of AI in clinics (Wang & Preininger, 2019; Alami *et al*., 2020). This underscores the importance of integrating risk perception theories with established technology acceptance models to capture the full spectrum of factors influencing decisions. This integrative view forms the theoretical framework presented in the following section.

Theoretical Framework

Value-Based Framework (UTAUT)

The Unified Theory of Acceptance and Use of Technology provides a robust theoretical foundation for the value-based analysis of AI adoption in this study, given its proven applicability across diverse healthcare technologies (Kalavani *et al*., 2018; Kim *et al*., 2016). When assessing the benefits, risks, usability, and overall utility of AI-enabled CDSSs, clinicians' value judgments are closely aligned

with UTAUT's core constructs: performance expectancy, effort expectancy, social influence, and facilitating conditions (Venkatesh *et al*., 2003).

Together, these dimensions explain how healthcare professionals perceive the value of AI technologies in improving practice performance and patient outcomes. Furthermore, the adequacy of training has been identified as a key enabler of technology adoption, as clinicians who receive sufficient instruction and institutional backing report higher confidence and lower resistance to AI systems (Ko *et al*., 2018; Alrahbi *et al*., 2021).

Accordingly, the framework offers a suitable lens for understanding how healthcare professionals translate perceived value and trust into actual adoption behaviour (Wills *et al*., 2008).

According to Adegboro *et al*. (2022), Perceived Benefits (PB) reflect the expectations of healthcare professionals that AI-based computer programs will improve clinical decision-making, patient care, diagnostic accuracy, and efficiency. Drawing on the UTAUT model, five key antecedents to Perceived Benefits are identified:

1. Performance Expectancy (PE): the belief that AI tools improve diagnostic accuracy and clinical efficiency (Mbelwa *et al*., 2019; Alabdullah *et al*., 2020);
2. Effort Expectancy (EE): the perceived ease of system use, interface intuitiveness, and integration into clinical workflow (Kohnke *et al*., 2014);
3. Social Influence (SI): the degree to which peers, superiors, and professional networks endorse the use of AI (Klingberg *et al*., 2019; Mbelwa *et al*., 2019);
4. Facilitating Conditions (FC): perceptions of support for AI usage from organisational infrastructure, technical resources and institutional policies (Orruño *et al*., 2011; Gagnon *et al*., 2012);
5. Training Adequacy (TA): the extent to which users feel adequately trained and confident in employing AI-based systems (Ko *et al*., 2018; Alrahbi *et al*., 2021).

Risk-Based Framework (Perceived Risk and Accountability Model)

This study also incorporates the risk-based framework proposed by Esmaeilzadeh (2020), which conceptualizes perceived risks as multifaceted constructs influenced by ethical, implementation-related, and technological concerns (Dwivedi *et al*., 2019). This model accounts for the uncertainty, accountability, and trust-related concerns that uniquely characterize AI applications in healthcare (Hengstler *et al*., 2016). In the context of AI-powered CDSSs, perceived risks are amplified by the high-stakes nature of medical decision making and clinicians' professional responsibility for patient outcomes (Hengstler *et al*., 2016). Common concerns include uncertainty about system reliability, lack of explainability, worrying about the security of data, and questions of liability when AI errors occur (Cath, 2018). While Esmaeilzadeh's framework was primarily applied to consumers and

administrative settings, it has been less explored among healthcare professionals, whose risk assessments are often influenced by professional accountability and patient safety considerations. Integrating UTAUT and risk perception frameworks therefore enables a balanced analytical approach that captures both the facilitators and barriers to AI-CDSSs adoption.

To ensure both conceptual clarity and measurement validity, key constructs were operationalized grounded in prior validated literature and classified into three primary areas of issues: *technological*, *ethical* and *implementation* concerns (Esmaeilzadeh, 2020). These domains are meant to act as antecedents to the Perceived Risks (PR) associated with AI-embedded CDSSs:

1. Technological Concerns include: 1) Perceived Performance Anxiety (PPA), the concern that AI tools may underperform or misinterpret clinical input as a result of data restrictions or algorithmic malfunction), and 2) Perceived Communication Barriers (PCB), namely complications in interpreting or integrating AI recommendations into human reasoning (Dwivedi *et al*., 2019; Lu & Schulz, 2024);
2. Ethical Concerns can be summarised in these constructs: 1) *Perceived Privacy Concerns* (PPC), including the sensitivity of healthcare data and potential breaches; 2) Perceived Mistrust in AI Mechanisms (PMT), which refer to the lack of confidence in the reliability and explainability of AI-embedded systems; and 3) Perceived Social Biases (PSB), reflecting concerns about embedded discrimination in algorithmic processes (Char *et al*., 2018; Reddy *et al*., 2020; Whittlestone *et al*., 2019);
3. Implementation-related Concerns cover: 1) Perceived Unregulated Standards (PUS), the perceived absence of clear guidelines governing AI adoption in healthcare, and 2) Perceived Liability Issues (PLI), referring to the uncertainty around responsibility in AI-induced clinical errors (Cath, 2018; Gupta & Kumari, 2017).

Hypotheses Development

Grounded in this synthesis, the proposed model conceptualizes perceived benefits and risks as key mediating constructs impacting healthcare professionals' behavioural intentions to adopt AI in medical practice (Dependent Variable, DV). This variable represents a reliable predictor of actual technology usage.

The model posits two primary pathways:

1. A *value-based* pathway, in which UTAUT constructs (performance expectancy, effort expectancy, social influence, facilitating conditions and training adequacy, Independent Variables, IVs) serve as antecedents of Perceived Benefits (Hp. 1.1-1.5), which are expected to positively influence clinicians' adoption intentions (Hp.1);

2. A *risk-based* pathway, in which technological, ethical, and implementation concerns (Independent Variables, IVs) act as antecedents of Perceived Risks (Hp2.1-Hp2.7), which in turn is hypothesized to negatively affect clinicians' intention to use AI in clinical practice (Hp.2) (Esmaeilzadeh *et al*., 2021).

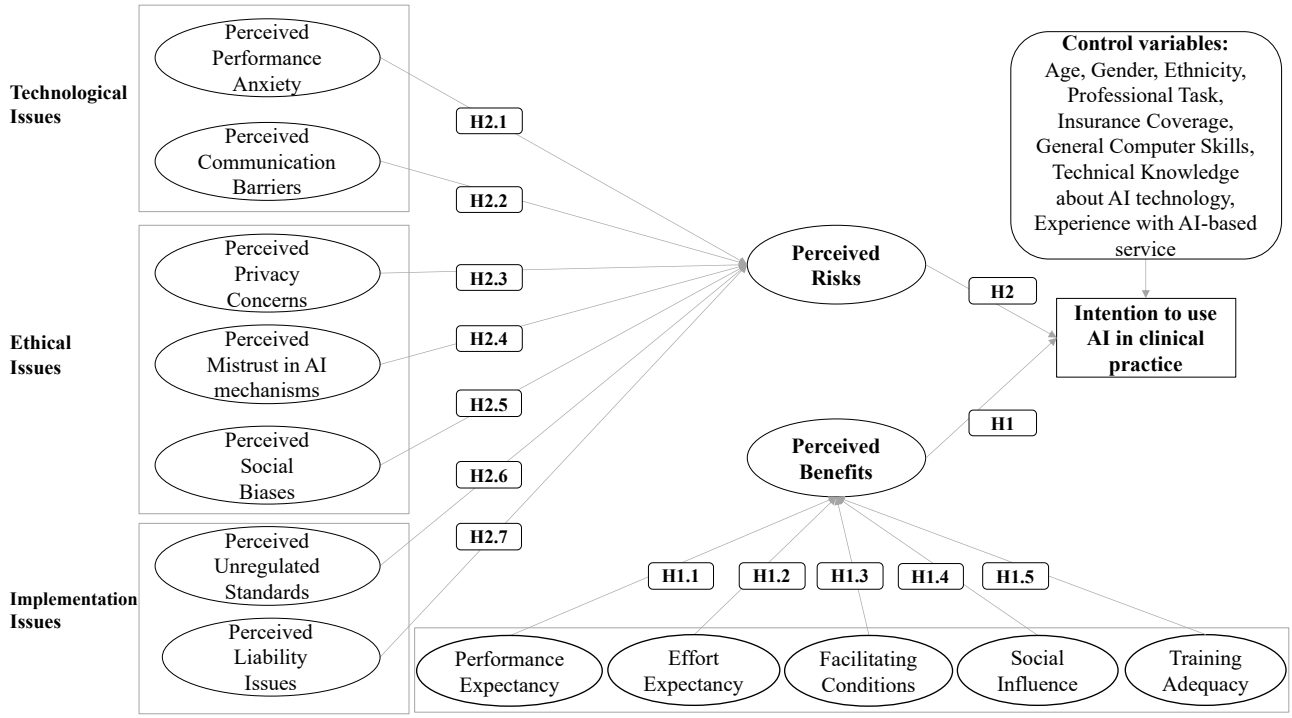
Table 7 – Hypotheses Development

	Research Question/Hypotheses	Sources	Description
	RQ1		How do healthcare professionals perceive the **benefits** of AI medical devices and, in particular, of clinical decision support systems (CDSS)?
	***Hp1*:** Perceived Benefits	Lo *et al*., 2018	Perceived *benefits* positively influence the intention to use AI-based tools
UTAUT constructs	***Hp1.1:*** Performance Expectancy	Mbelwa *et al*., 2019; Alabdullah *et al*., 2020; Kohnke *et al*., 2014	Higher *performance expectancy* is positively associated with perceived benefits of AI-based tools
	Hp1.2: Effort Expectancy	Mbelwa *et al*., 2019; Alabdullah *et al*., 2020; Kohnke *et al*., 2014	Greater *effort expectancy* is positively associated with perceived benefits, as ease of use enhances perceived value
	Hp1.3: Facilitating Conditions	Orruño *et al*., 2011; Gagnon *et al*.; Klingberg *et al*., 2019; Alabdullah *et al*., 2020; Mbelwa *et al*., 2019; Kohnke *et al*., 2014	Perceived availability of necessary resources and support (*facilitating conditions*) positively influences perceived benefits
	Hp1.4: Social Influence	Klingberg *et al*., 2019; Alabdullah *et al*., 2020; Mbelwa *et al*., 2019; Kohnke *et al*., 2014	Greater *social influence* from colleagues or leadership increases the perceived benefits of AI-based clinical tools
	Hp1.5: Training Adequacy	Mbelwa *et al*., 2019	Greater *training adequacy* positively influences the perceived benefits of AI-based tools
	RQ2		What are the perceived **risks** associated with the adoption of AI medical devices among healthcare providers?
	Hp2: Perceived Risks	Bansal *et al*., 2016	Higher perceived *risks* negatively affect the intention to use AI-based tools
Technological Concerns	***Hp2.1***: Perceived Performance Anxiety	Sarin *et al.*, 2003	The greater the perceived *performance anxiety*, the higher the perceived risks
	Hp2.2: Perceived Communication Barriers	Lu *et al*., 2019	The presence of *communication barriers* due to AI usage positively correlates with increased perceived risks
Ethical Concerns	***Hp2.3***: Perceived Privacy Concerns	Stewart and Segars, 2002	Higher *perceived privacy concerns* lead to increased perceived risks
	Hp2.4: Perceived Mistrust in AI Mechanisms	Luxton, 2019	Perceived *mistrust* in AI mechanisms is positively correlated with perceived risks
	Hp2.5: Perceived Social Biases	Reddy *et al*., 2020	Perceived *social biases* in AI positively influence perceived risks
Implementation Concerns	***Hp2.6***: Perceived Unregulated Standards	Cath, 2018	Perceived *unregulated standards* are positively associated with perceived risks
	Hp2.7: Perceived Liability Issues	Laï *et al*., 2020	Perceived *liability issues* positively correlate with increased perceived risks

Source: our elaboration

The model also comprises numerous control variables validated in prior research, including age, gender, clinical expertise, insurance coverage, current professional activity, technical familiarity with AI technology, and previous experience with AI-based services (Esmaeilzadeh *et al*., 2021).

Figure 12 - Research Model



From Conceptual Framework to Measurement Model

To empirically test the proposed framework, the study's measurement model was developed through a systematic process to ensure both theoretical alignment and contextual validity. This entailed four sequential steps: 1) establishing the base framework; 2) integrating complementary constructs; 3) adapting UTAUT determinants, and 4) validating item relevance through pre-testing.

The initial structure was based on Esmaeilzadeh's (2020) model of AI adoption, which identifies key dimensions of perceived risk (technological, ethical and implementation-related). All constructs and items from this framework were adapted to target healthcare professional context, rather than consumers.

Pre-validated items from earlier research in telemedicine, mobile health, and teledermatology (e.g., Hajesmaeel Gohari & Bahaadinbeigy, 2021; Hailu *et al*., 2025; Tao *et al*., 2025) were initially taken into consideration for the behavioural intention and training constructs during the scale development phase. However, these items were ultimately removed from the completed survey for the following reasons: domain misalignment, semantic redundancy, lack of contextual relevance, survey burden considerations.

All constructs were measured using reflective indicators on a 7-point Likert scale ranging from 1 (*Strongly Disagree*) to 7 (*Strongly Agree*). The final instrument comprised 30 items related to Esmaeilzadeh's constructs and 25 items from UTAUT, plus demographic and background questions (e.g., age, gender, clinical specialty, familiarity with AI, and previous AI usage).

To tailor the instrument for a clinical audience, all items were reworded to align with healthcare terminology. For instance, general terms such as "*system*" were replaced with "*AI-based Clinical Decision Support Systems*", and item wording was tailored to reflect realistic scenarios relevant to clinical workflows, such as diagnostics, treatment planning, patient safety, and communication.

A pre-testing phase involving 5 healthcare professionals was conducted to assess face validity, clarity, and relevance. Feedback from this pilot informed minor rewordings, particularly to ensure clinical terminology was contextually appropriate and unambiguous.

Methodology

In this study, a cross-sectional survey design was employed, utilizing a structured questionnaire adapted from previously validated instruments in the domains of healthcare technology adoption and AI risk perception (Venkatesh *et al*., 2003; Esmaeizadeh, 2020; Petkus *et al*., 2020; Roudi *et al*., 2022).

The target population consisted of licensed healthcare professionals across different roles and specialties, including physicians, surgeons, nurses, radiographers, technicians and allied health professionals. The sampling strategy followed a non-probability purposive approach, utilized to access participants from professional mailing lists, clinical networks, and online forums. The eligibility criteria required the respondents to be: 1) aged 18 years or above; 2) currently practicing clinically; 3) fluent in English; and 4) professionally qualified healthcare workers. Participation was voluntary and anonymous. Electronic informed consent was obtained prior to survey commencement. Meetings to provide institutional contact information were offered, with an invitation for participants to contact the authors for any questions. Ethical approval was granted by the Faculty Ethics Committee (FEC) of the University of Southampton [ERGO No.: 103435]. The ideal sample size calculation, which is crucial for attaining the reliability and validity of findings, was derived using the formula:

$$n = \frac{Z^2 PQ}{d^2}$$

Where, n = required sample size; Z = precision level (1.96 for 95% confidence level); P = prevalence estimated value from previous studies; Q = 1-P. Drawing on prior research focused on evaluating AI adoption in healthcare organizations where the relative adoption and perceived usefulness of AI-CDSSs range between 70% and 80% (Dingel *et al*., 2024; Scott *et al*., 2024; Zheng *et al*., 2025), the present study estimated that about 76% of healthcare professionals will consider AI adoption as relevant in their practice. This assumption was employed as the baseline proportion (P) to determine the minimum sample size that would ensure reliable and representative outcomes.

Considering the exploratory nature of the study and to provide feasibility in different healthcare contexts, a 7% margin of error was employed (d= 0.05). This decision was grounded in a compromise between the statistical accuracy and the recruitment logistics, while still providing precise group-level estimates. Consistent with the above-mentioned configuration, the minimum required sample size was approximated to be around 197 participants (Sharma *et al*., 2009; Althubaiti, 2022; Banerjee *et al*., 2024).

Data Analysis

The survey was administered online using Qualtrics[XM9]. Participants accessed the questionnaire through a secure, anonymized link. The instrument was open for responses for 14 weeks, and reminder emails were sent to improve response rates. Participants were required to complete the survey in a single session; partial responses were excluded. No personally identifying information was collected. Respondents were provided with a brief, accessible definition of AI and AI-based CDSSs to ensure shared understanding.

For the Italian sample, the back-translation method was employed, in accordance with the recommended best practices from Bujang *et al*. (2022), which provide a "*Step-by-Step Guideline to Questionnaire Validation Research*". The process involved two groups: one consisting of a subject-matter expert and the other including an English linguistic expert, who conducted the initial translation. A different subject-matter expert and another linguistic expert then performed the back-translation. The researcher examined all four versions before finalizing the pre-final version of the questionnaire.

A total of 215 individuals completed the survey (incomplete answers were discarded). This exceeded the minimum sample size required for the model structure. Consequently, the study is adequately powered, with a number of samples potentially sufficient to reduce sampling errors and minimise Type II errors. The distribution across Italy and the UK was highly uneven, with 202 Italian (94%) and 13 UK (6%) respondents. Therefore, cross-national statistical comparisons were not conducted, and analyses were performed on the combined dataset.

Before testing the measurement model, the dataset was screened for normality, given its importance in ensuring unbiased parameter estimates and valid inferences (Hair *et al*., 2006). Normality was screened by exploring *skewness* and *kurtosis* statistics of all observed indicators of latent constructs in the model. The absolute values of 2 for skewness and 7 for kurtosis indicate the threshold, below which deviation from normality is acceptable (West *et al*., 1995). The analysis produced skewness values ranging from -1.007 to -0.037 and kurtosis values ranging from -0.983 to

[9] Qualtrics[XM]. https://www.qualtrics.com/

1.278 for all the items. All the values were within the prescribed ranges, so no deviation from normality was noted in the distribution of data and it was suitable for further Confirmatory Factor Analysis (CFA) and Structural Equation Modelling (SEM).

Following the cutoff points recommended by Gefen *et al*. (2000) and Hair *et al*. (2011), convergent validity was evaluated by examining standardized factor loadings, composite reliability (CR), and average variance extracted (AVE) (Table 8). Standardized loadings of ≥ 0.70, CR ≥ 0.70, and AVE ≥ 0.50 indicate acceptable convergent validity.

According to the results of the confirmatory factor analysis (CFA), most constructs met these thresholds. Significant reliability and convergent validity were demonstrated by Effort Expectancy (EE) (CR = 0.919, AVE = 0.590), Intention to Use (INT) (CR = 0.931, AVE = 0.729), Perceived Communication Barriers (PCB) (CR = 0.941, AVE = 0.764), Perceived Unregulated Standards (PUS) (CR = 0.943, AVE = 0.767), and Perceived Performance Anxiety (PPA) (CR = 0.868, AVE = 0.569).

However, Facilitating Conditions (FC) did not meet the recommended AVE threshold (CR = 0.703, AVE = 0.312), largely due to four items, FC4 (λ = 0.446), FC5 (λ = 0.324), FC7 (λ = 0.524), and FC8 (λ = 0.271), with loadings significantly below 0.70. Similarly, one item in Perceived Social Biases (PSB) showed a low loading (PSB1, λ = 0.499), although the overall AVE (0.573) and CR (0.866) were still above the threshold. Despite including an item below the optimal loading level (TA3, λ = 0.646), Training Adequacy (TA) achieved the lowest allowable CR value (0.700) and an AVE of 0.542.

The above-mentioned thresholds were met or exceeded by all other constructs, including Perceived Risks (PR) (CR = 0.902, AVE = 0.570), Performance Expectancy (PE) (CR = 0.878, AVE = 0.592), Perceived Privacy Concerns (PPC) (CR = 0.925, AVE = 0.679), Perceived Mistrust (PMT) (CR = 0.859, AVE = 0.553), and Social Influence (SI) (CR = 0.878, AVE = 0.596).

Lastly, to increase construct reliability, low-loading items from FC were eliminated, but PSB was kept in its original configuration because of its respectable overall metrics.

Table 8 - Results of convergent validity

Construct	Item	Standardized Factor Loading (> 0.7)	Composite Reliability (> 0.7)	AVE (> 0.5)
Perceived Benefits	PB1	0,747	0,946	0,597
	PB10	0,732		
	PB11	0,736		
	PB12	0,857		
	PB2	0,786		
	PB3	0,728		
	PB4	0,86		
	PB5	0,815		
	PB6	0,653		
	PB7	0,809		
	PB8	0,784		

	PB9	0,771		
Performance Expectancy	PE1	0,752	0,88	0,588
	PE2	0,864		
	PE3	0,709		
	PE4	0,732		
	PE5	0,782		
Effort Expectancy	EE1	0,719	0,893	0,582
	EE2	0,883		
	EE3	0,903		
	EE4	0,866		
	EE5	0,656		
	EE6	0,729		
	EE7	0,672		
	EE8	0,664		
Facilitating Conditions	FC2	0,878	0,797	0,665
	FC3	0,744		
Social Influence	SI1	0,559	0,876	0,604
	SI2	0,8		
	SI3	0,901		
	SI4	0,86		
	SI5	0,69		
Training Adequacy	TA1	0,831	0,708	0,554
	TA3	0,634		
Perceived Risks	PRisk1	0,72	0,899	0,554
	PRisk2	0,828		
	PRisk3	0,739		
	PRisk4	0,799		
	PRisk5	0,839		
	PRisk6	0,727		
	PRisk7	0,606		
Perceived Performance Anxiety	PPA1	0,678	0,864	0,566
	PPA2	0,839		
	PPA3	0,711		
	PPA4	0,77		
	PPA5	0,763		
Perceived Communication Barriers	PCB1	0,901	0,939	0,769
	PCB2	0,885		
	PCB3	0,96		
	PCB4	0,965		
	PCB5	0,612		
Perceived Privacy Concerns	PPC1	0,533	0,922	0,686
	PPC2	0,771		
	PPC3	0,925		
	PPC4	0,916		
	PPC5	0,865		
	PPC6	0,869		
Perceived Mistrust	PMis1	0,649	0,852	0,543
	PMis2	0,71		
	PMis3	0,898		
	PMis4	0,795		
	PMis5	0,633		
Perceived Social Biases	PSB1	0,499	0,883	0,593
	PSB2	0,841		
	PSB3	0,83		
	PSB4	0,702		
	PSB5	0,852		
Perceived Unregulated Standards	PUS1	0,811	0,939	0,763
	PUS2	0,857		
	PUS3	0,885		
	PUS4	0,906		
	PUS5	0,914		
Perceived Liability Issues	PLI1	0,861	0,932	0,697
	PLI2	0,839		

	PLI3	0,898		
	PLI4	0,891		
	PLI5	0,843		
	PLI6	0,666		
Intention to Use	IntUSE1	0,882	0,932	0,727
	IntUSE2	0,867		
	IntUSE3	0,902		
	IntUSE4	0,869		
	IntUSE5	0,741		

Source: our elaboration

Discriminant validity was evaluated using the Fornell-Larcker criterion, by comparing the inter-construct correlations (off-diagonal values) to the square root of the Average Variance Extracted (AVE) for each construct (shown on the diagonal). Table 9 demonstrates that each construct's square root of AVE was higher than its correlations with every other construct, demonstrating satisfactory discriminant validity (Fornell & Larcker, 1981). This result demonstrates that every latent variable in the measurement model is empirically different from the others.

Table 9 – Results of Discriminant Validity

Construct	Mean	SD	EE	FC	INT	PB	PCB	PE	PLI	PMT	PPA	PPC	PR	PSB	PUS	SI	TA
EE	4,39	1,1															
FC	5,04	1,23	0,274														
INT	4,86	1,19	0,403	0,591													
PB	4,79	1,09	0,46	0,539	0,688												
PCB	4,63	1,53	-0,155	-0,053	-0,225	-0,179											
PE	4,63	1,18	0,549	0,418	0,574	0,697	-0,196										
PLI	4,99	1,31	-0,164	0,046	-0,087	-0,135	0,461	-0,19									
PMT	4,33	1,04	0,303	0,291	0,423	0,416	-0,052	0,373	-0,04								
PPA	4,42	1,16	-0,152	-0,134	-0,302	-0,26	0,286	-0,246	0,338	-0,339							
PPC	4,36	1,42	-0,121	-0,071	-0,185	-0,216	0,364	-0,18	0,432	-0,118	0,407						
PR	4,4	1,04	-0,207	-0,177	-0,36	-0,309	0,496	-0,283	0,565	-0,214	0,524	0,494					
PSB	4,17	1,23	-0,165	-0,077	-0,204	-0,186	0,335	-0,218	0,339	-0,233	0,573	0,492	0,44				
PUS	5,05	1,22	-0,017	0,121	-0,048	-0,089	0,4	-0,104	0,501	-0,069	0,399	0,455	0,479	0,42			
SI	4,08	1,13	0,345	0,283	0,313	0,456	-0,004	0,382	-0,032	0,183	-0,129	-0,047	-0,173	-0,085	-0,026	0,777	
TA	5,11	1,23	0,144	0,46	0,507	0,412	-0,023	0,336	0,06	0,136	0,023	-0,119	-0,042	0,044	0,186	0,114	0,744

Source: our elaboration

Although the correlations between the constructs were not evident, we assessed multicollinearity by calculating the Variance Inflation Factor (VIF) and tolerance values for the predictor variables. The resulting tolerance values ranged from 0.37 to 0.74, surpassing the minimum threshold of 0.1, and the VIF values ranged from 1.36 to 2.70, which were significantly below the suggested cutoff value of 5 (Hair *et al*., 2011). These findings suggest that multicollinearity was not an issue for this investigation.

All measurement items were subjected to an unrotated principal axis factor analysis to evaluate the possibility of common method bias using Harman's single-factor test (Podsakoff *et al*., 2003). According to the findings, the first factor explained 23.55% of the variance overall, far less than the

50% threshold that is advised. This suggests that common method variance is not likely to pose a significant risk to the current investigation.

Results

Descriptive Statistics

Table 10 outlines the characteristics of the respondents. Demographic data show that most participants were female (57.7%), while males represented 41.4% of the sample. One respondent chose not to reveal their gender, while another identified as third gender or non-binary. Over half of respondents (51.1%) were 50 years of age or older. The remaining percentages were split more evenly among younger age groups: 13.5% were 40–49, 20.5% were 30–39, and 14.9% were 20–29. The largest groups in terms of professional qualifications were medical doctors (34.6%) and nurses (39.3%), followed by nurse auxiliaries (12.6%), technicians (4.7%), and other professions (8.9%). A wide range of clinical specialties were represented: radiology (14.7%) and primary care (13.3%) were the most reported specialties, followed by emergency medicine (12.0%) and dermatology (6.7%). Other areas of specializations also appeared, such as paediatric gastroenterology, dentistry, ophthalmology, sports medicine, neurology, pathology, oncology, and obstetrics and gynaecology.

Regarding the insurance coverage, the most prevalent type of professional indemnity insurance was self-funded (55.8%), immediately followed by employer-provided professional indemnity insurance (38.1%). Other forms, such as mandatory malpractice insurance (28.4%) and vehicle insurance (20.9%) appeared to be less common. Only 6.0% of respondents mentioned coverage types like property insurance and personal accident insurance.

Experience levels differed greatly: about 30.2% had over thirty years of clinical experience, 24.7% had five years or less, and the rest declared evenly distributed clinical experience from 6 to 30 years.

Regarding the use of AI, 64.2% reported using AI applications for non-healthcare purposes (such as smart devices or financial decision-making) and 60% of respondents had employed AI in clinical settings. Only 19.6% of respondents felt very or extremely familiar with AI, compared to 25.1% who felt only slightly familiar and completely unfamiliar (9.8%).

The experiences of those who had used clinical AI were not all the same: 38.0% found it neither easy nor difficult to use, whereas a total of 26.4% characterized the experience as very or extremely challenging. As for the perceived impact on workload, 57.4% reported surprisingly no change, 27.9% reported some reduction and only 14.7% reported an increase.

Once more, most answers centred on neutrality when asked about learning to use AI: 38.8% said it was neither easy nor difficult. Only 27.2% thought it was somewhat to extremely easy, while 34.1% thought it was moderately to extremely difficult. Among respondents who effectively used AI systems

in their clinical practice, there was also variation in the level of understanding of the outputs: 49.6% reported moderate to full understanding and only 19.4% reported little to no understanding.

Finally, when asked about common clinical tasks that involve the use of CDSS, the most mentioned uses were diagnosis assistance (32.6%), interpretation of investigation results (18.6%) and drug dosage calculations (14.9%). Several participants discussed specific tasks where CDSS tools are used in addition to the typical ones, including research and evidence gathering, psychological support, decision-support algorithms, translation support, and patient-specific care planning.

Table 10 – Sample Characteristics

Variable	Key Categories
Gender	Male (41.4%) Female (57.7%)
Age	20–39 (35.4%) 40–59 (39.5%) 60+ (25.1%)
Professional Qualification	Medical Doctor (34.6%) Nurse (39.3%) Others (26.1%)
Clinical Experience	≤10 years (35.4%) 11–30 years (34.4%) >30 years (30.2%)
Used AI in Healthcare	Yes (60%) No (40%)
Ease of Use (AI healthcare)	Neutral (38%) Challenging (44.2%) Easy (17.9%)
Understanding AI Output	Well–Very well (37.3%) Somewhat (23.3%) Little/None (19.4%)

Source: our elaboration

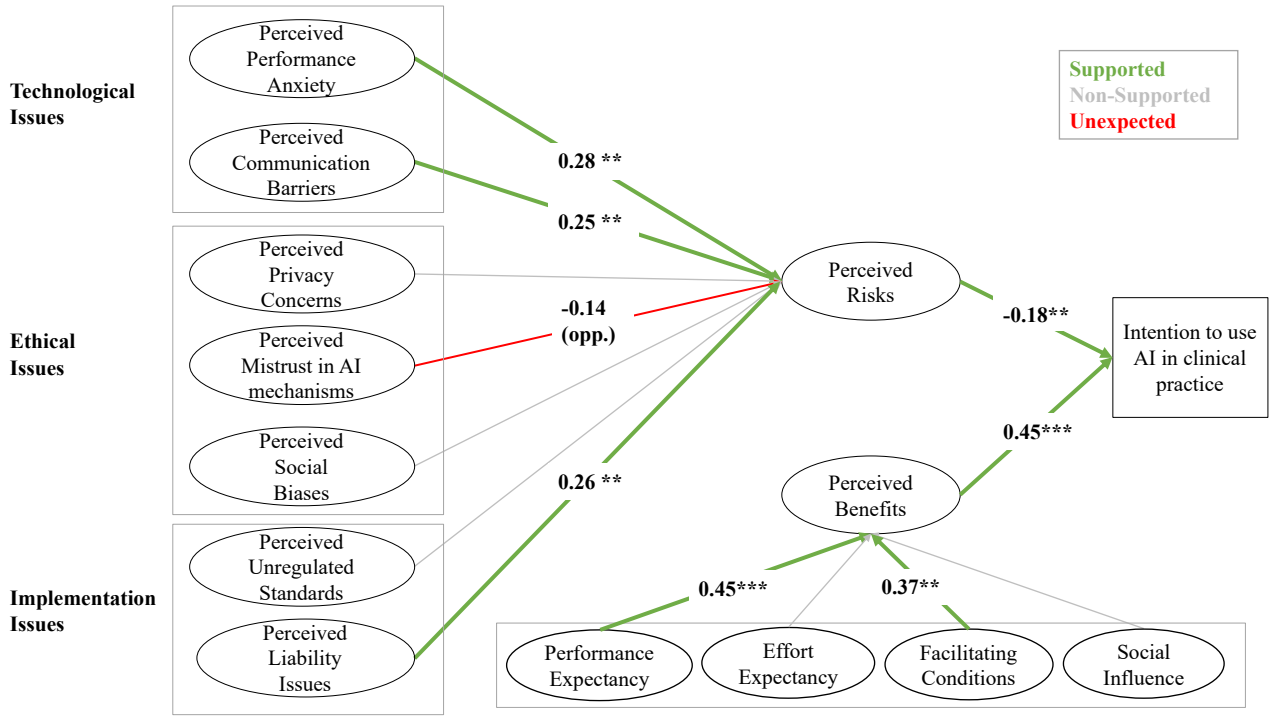
Structural Model

The structural model was estimated using MLR (Robust Maximum Likelihood with Yuan–Bentler T2* correction). The model converged successfully with 183 free parameters and demonstrated an acceptable overall fit. Fit indices showed a χ^2/df ratio of roughly 1.59 and a scaled chi-square of χ^2 (944) = 1501, p <.001. Adequacy was further supported by the robust fit indices, which were RMSEA =.058 (90% CI [.052–.063]), TLI =.902, and CFI =.911. Collectively, these values suggest that the hypothesised model offers a plausible representation of the observed data (Figure 13).

Regarding UTAUT sub-constructs, both Performance Expectancy (β =.45, p <.001) and Facilitating Conditions (β =.37, p <.001) significantly increased intention to use AI, supporting hypotheses Hp1.1 (performance expectancy → INT) and Hp1.3 (facilitating conditions → INT). Conversely, Hp1.2 (effort expectancy → INT) and Hp1.4 (social influence → INT) were not

supported, as these factors did not exhibit discernible associations with intention to use AI. Overall, the model had a moderate but significant explanatory power in predicting clinicians' willingness to use AI-enabled CDSSs, explaining 46% of the variance in INT (R^2 =.46).

Figure 13 – Model Paths



Source: our elaboration

The results partially validate the proposed antecedents of perceived risk (PR). Perceived Performance Anxiety (PPA, Hp2.1) showed a significant positive impact on PR (β =.28, p =.005), as well as Perceived Communication Barriers (PCB, Hp2.2) (β =.25, p =.002). Furthermore, Perceived Liability Issues (PLI, Hp2.7) emerged as a third strong and significant predictor of perceived risks (β = .26, p = .002). On the other hand, Perceived Mistrust (PMT, Hp2.4) showed a negative impact on PR (β = –.14, p =.040), in contrast with the original hypothesis, which predicted a positive relationship between mistrust and perceived risks. Hp2.3 (privacy concerns → PR), Hp2.5 (social bias → PR), and Hp2.6 (unregulated standards → PR) were not supported, since these constructs had no discernible effect on risk perceptions.

The findings support the value-based approach to technology adoption by highlighting the influence of perceived risks and benefits on clinicians' adoption intentions. Strong support was found for Hp1, which proposed that intention to use AI-based CDSSs would be positively impacted by Perceived Benefits (β =.45, p <.001). Hp2, which predicted that intention would be negatively impacted by Perceived Risks, was also validated (β = -.18, p =.002).

Table 11 – Results of hypotheses testing

Hypothesis	Path Tested	St. Coeff.	P-value	Outcome
Hp1	PB → INT	β = .45***	p < .001	Supported
Hp1.1	PE → PB/INT	β = .45***	p < .001	Supported
Hp1.2	EE → PB/INT	n.s.		Not Supported
Hp1.3	FC → PB/INT	β = .37***	p < .001	Supported
Hp1.4	SI → PB/INT	n.s.		Not Supported
Hp1.5	Training Adequacy → PB	(not tested in dataset)		–
Hp2	PR → INT	β = –.18**	p = .002	Supported
Hp2.1	PPA → PR	β = .28**	p = .005	Supported
Hp2.2	PCB → PR	β = .25**	p = .002	Supported
Hp2.3	PPC → PR	n.s.		Not Supported
Hp2.4	PMT → PR	β = –.14*	p = .040	Not Supported (opposite direction)
Hp2.5	PSB → PR	n.s.		Not Supported
Hp2.6	PUS → PR	n.s.		Not Supported
Hp2.7	PLI → PR	β = .26**	p = .002	Supported

Source: our elaboration

Discussion

The present study investigated the factors that influence clinicians' intentions to use AI-based clinical decision support systems (CDSSs) using a value-based risk-benefit framework. Although the results are generally in line with earlier studies (Esmaeilzadeh, 2020), which highlighted the predictive value of perceived risks and benefits over conventional technology acceptance constructs, significant differences were revealed as a result of our sample's professional makeup and contextual setting.

Consistently with prior research, our findings confirm that adoption intentions are more strongly influenced by Perceived Benefits than by Perceived Risks. Clinicians' perception of benefits was greatly enhanced by Performance Expectancy (PE) and Facilitating Conditions (FC), and this in turn strongly predicted their intention to use AI-enabled CDSSs. On the other hand, Social Influence (SI) and Effort Expectancy (EE) were not significant.

These results suggest that institutional support and functional improvement are more relevant drivers of adoption of technology by healthcare professionals than peer influence or ease of use. This supports the notion that emphasizing clinical utility and infrastructural readiness is important for adoption. It also aligns with Esmaeilzadeh's assertion that value-based evaluations are vital for the adoption of AI tools in healthcare. In terms of risk perceptions, there were several similarities and differences with the study published by Esmaeilzadeh (2020). Both studies found that risk beliefs are significantly influenced by technological issues. In our sample, Perceived Performance Anxiety (PPA) and Perceived Communication Barriers (PCB) considerably raised risk perceptions, supporting previous findings that uncertainties about reliability and the decline of interpersonal communication are persistent issues. Similarly, Perceived Liability Issues (PLI), which represent physicians' responsibility for patient safety and potential legal repercussions, were also a powerful predictor of risks. These findings support Esmaeilzadeh's conclusion that implementation concerns are important,

but they also imply that liability risks are more significant for practicing clinicians than worries about unregulated standards (PUS), which were not significant in this study. One potential explanation for this phenomenon is that medical professionals operate under the assumption that AI systems employed within the healthcare sector will receive regulatory approval prior to their implementation. This assumption would serve to eliminate any perceived ambiguity regarding standards, yet it would not guarantee accountability in the event of errors occurring.

However, in contrast to the findings of earlier research, the ethical issues in this study manifested differently. Esmaeilzadeh (2020) discovered that while Privacy Concerns (PPC) and Social Biases (PSB) were not significant, Perceived Mistrust (PMT) of AI mechanisms raised risk perceptions. Conversely, our results showed that while PMT unexpectedly decreased perceived risks, Perceived Privacy Concerns and Perceived Social Biases did not predict risks. This unexpected finding might be due to a contextual difference: clinicians in our sample, working in structured healthcare systems in Italy and the UK, might believe that institutional safeguards, professional oversight, and human-in-the-loop methods help to reduce mistrust in AI. In this sense, mistrust may not increase perceived risks but rather serve as a stand-in for professional scepticism, which allows clinicians to evaluate AI recommendations critically without necessarily viewing them as dangerous. This phenomenon stands in contrast to that observed among lay users, who, as demonstrated by Esmaeilzadeh, tend to associate mistrust with heightened risk perceptions due to their unawareness of clinical and regulatory checks.

Implementation Concerns are the subject of another distinction. In the model under consideration, it appeared that only Perceived Liability Issues (PLI) were significant, even though Esmaeilzadeh reported both PLI and Perceived Unregulated Standards (PUS) as significant predictors of risks. Once more, variations in sample composition could be the cause of this. While clinicians are used to strict oversight procedures (such as EMA, MHRA, or NICE approvals), lay participants might not be confident in the adequacy of regulatory safeguards. As a result, they perceive liability – rather than standards – as the most significant regulatory risk.

Building upon these considerations, clinicians' adoption intentions appear to be contingent less on technical performance and more on perceived alignment with ethical and professional values. This finding suggests that value congruence – that is, the alignment between clinicians' professional norms and the perceived logic of AI – may serve as a more robust predictor of adoption than traditional utilitarian factors.

In conclusion, the findings of both studies indicated the presence of a dual influence of risks and benefits on adoption intentions, albeit with divergent weights attributed to each. The present findings lend support to Esmaeilzadeh's assertion that perceived benefits have a greater positive impact than perceived risks (negative impact). This lends further support to the notion that the presentation of AI

in terms of clinical utility and decision support, as opposed to the attempt to mitigate concerns, is more likely to encourage its adoption.

Conclusions

The present study demonstrates that, within clinical settings, decisions on whether to use new AI-based technologies are less influenced by peer endorsement or ease of use and more by the perceived utility, accountability, and ethical compliance of the AI system. Notably, the inverse association between Perceived Mistrust and risk perception further suggests that clinicians' scepticism does not necessary translate into rejection of AI; instead, it may serve as a protective and clinical mechanism that facilitates safer and more informed use.

From a theoretical standpoint, this work extends the applicability of value-based models to professional settings, identifying value congruence i.e., the alignment between clinicians' ethical and professional norms and the perceived logic of AI, as a novel and potentially decisive predictor of adoption. On a practical level, the findings suggest that AI integration strategies should prioritise clear accountability mechanisms, strong institutional support, and targeted professional training to foster trust-building and confidence in AI-assisted decision-making.

Future research should explore these dynamics across more diverse healthcare systems, cultural contexts, and medical specialties to assess the generalisability of the model. Comparative and longitudinal studies may further clarify how evolving clinical experience with AI shapes perceptions of risks, benefits, and ethical fit over time.

References

Adegboro, C. O., Choudhury, A., Asan, O., & Kelly, M. M. (2022). Artificial Intelligence to Improve Health Outcomes in the NICU and PICU: A Systematic Review. *Hospital pediatrics*, *12*(1), 93–110. https://doi.org/10.1542/hpeds.2021-006094

Alabdullah, J., Lunen, B., Claiborne, D., Daniel, S., Yen, C.-J., & Gustin, T. (2020). Application of the Unified Theory of Acceptance and Use of Technology model to predict dental students' behavioral intention to use teledentistry. *Journal of Dental Education, 84*(8), 909–917. https://doi.org/10.1002/jdd.12304

Alami, H., Lehoux, P., Auclair, Y., de Guise, M., Gagnon, M. P., Shaw, J., Roy, D., Fleet, R., Ag Ahmed, M. A., & Fortin, J. P. (2020). Artificial Intelligence and Health Technology Assessment: Anticipating a New Level of Complexity. *Journal of medical Internet research*, *22*(7), e17707. https://doi.org/10.2196/17707

Alrahbi, D.A., Khan, M., & Hussain, M. (2019). Exploring the motivators of technology adoption in healthcare. *International Journal of Healthcare Management, 14*, 50 - 63. https://doi.org/10.1080/20479700.2019.1607451

Althubaiti A. (2022). Sample size determination: A practical guide for health researchers. *Journal of general and family medicine*, *24*(2), 72–78. https://doi.org/10.1002/jgf2.600

Asan, O., Bayrak, A.E., Choudhury, A. (2020). Artificial intelligence and human trust in healthcare: focus on clinicians. *Journal of Medical Internet Research*, 22 (6), e15154. https://doi.org/10.2196/15154

Asan, O., Choudhury, A. (2021). Research trends in Artificial Intelligence applications in human factors health care: mapping review. *JMIR Human Factors*, 8(2), e28236. https://doi.org/10.2196/28236

Banerjee, A., Sarangi, P. K., & Kumar, S. (2024). Medical Doctors' Perceptions of Artificial Intelligence (AI) in Healthcare. *Cureus*, *16*(9), e70508. https://doi.org/10.7759/cureus.70508

Bansal, G., Zahedi, F. M., & Gefen, D. (2016). Do context and personality matter? Trust and privacy concerns in disclosing private information online. *Information & Management, 53*(1), 1–21. https://doi.org/10.1016/j.im.2015.08.001

Broadbent, E., Kahokehr, A., Booth, R. J., Thomas, J., Windsor, J. A., Buchanan, C. M., Wheeler, B. R., Sammour, T., & Hill, A. G. (2012). A brief relaxation intervention reduces stress and improves surgical wound healing response: a randomised trial. *Brain, behavior, and immunity*, *26*(2), 212–217. https://doi.org/10.1016/j.bbi.2011.06.014

Bujang, M. A., Hon, Y. K., Lee, K., & Yee, K. (2022). *A Step-by-step Guide to Questionnaire Validation Research*. https://doi.org/10.5281/zenodo.6801209

Camilleri, M.A. (2024). Artificial intelligence governance: ethical considerations and implications for social responsibility. *Expert Systems*, *41*(7). https://doi.org/10.1111/exsy.13406

Cath C. (2018). Governing artificial intelligence: ethical, legal and technical opportunities and challenges. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, *376*(2133), 20180080. https://doi.org/10.1098/rsta.2018.0080

Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing Machine Learning in Health Care - Addressing Ethical Challenges. *The New England journal of medicine*, *378*(11), 981–983. https://doi.org/10.1056/NEJMp1714229

Choudhury A. (2022). Factors influencing clinicians' willingness to use an AI-based clinical decision support system. *Frontiers in digital health*, *4*, 920662. https://doi.org/10.3389/fdgth.2022.920662

Choudhury, A., & Shamszare, H. (2023). Investigating the Impact of User Trust on the Adoption and Use of ChatGPT: Survey Analysis. *Journal of medical Internet research*, *25*, e47184. https://doi.org/10.2196/47184

Cimperman, M., Makovec Brenčič, M., & Trkman, P. (2016). Analyzing older users' home telehealth services acceptance behavior-applying an Extended UTAUT model. *International journal of medical informatics*, *90*, 22–31. https://doi.org/10.1016/j.ijmedinf.2016.03.002

Collado-Mesa, F., Alvarez, E., & Arheart, K. (2018). The Role of Artificial Intelligence in Diagnostic Radiology: A Survey at a Single Radiology Residency Training Program. *Journal of the American College of Radiology : JACR*, *15*(12), 1753–1757. https://doi.org/10.1016/j.jacr.2017.12.021

Davis, F. (1989) Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly, 13, 319-340. https://doi.org/10.2307/249008

Dingel, J., Kleine, A. K., Cecil, J., Sigl, A. L., Lermer, E., & Gaube, S. (2024). Predictors of Health Care Practitioners' Intention to Use AI-Enabled Clinical Decision Support Systems: Meta-Analysis Based on the Unified Theory of Acceptance and Use of Technology. *Journal of medical Internet research*, *26*, e57224. https://doi.org/10.2196/57224

Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2019). Re-examining the Unified Theory of Acceptance and Use of Technology (UTAUT): Towards a Revised Theoretical Model. *Information Systems Frontiers*, *21*(3), 719-734. https://doi.org/10.1007/s10796-017-9774-y

Elhaddad, M., & Hamam, S. (2024). AI-Driven Clinical Decision Support Systems: An Ongoing Pursuit of Potential. *Cureus*, *16*(4), e57728. https://doi.org/10.7759/cureus.57728

Esmaeilzadeh, P. (2020). Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. *BMC Medical Informatics and Decision Making*, *20*(1), 170. https://doi.org/10.1186/s12911-020-01191-1

Esmaeilzadeh, P., Mirzaei, T., & Dharanikota, S. (2021). Patients' Perceptions Toward Human-Artificial Intelligence Interaction in Health Care: Experimental Study. *Journal of medical Internet research*, *23*(11), e25856. https://doi.org/10.2196/25856

Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S. and Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, *25*(1), 24-29. https://doi.org/10.1038/s41591-018-0316-z

European Society of Radiology (ESR) (2019). Impact of artificial intelligence on radiology: a EuroAIM survey among members of the European Society of Radiology. *Insights into imaging*, *10*(1), 105. https://doi.org/10.1186/s13244-019-0798-3

Fornell, C., & Larcker, D. F. (1981). Structural Equation Models with Unobservable Variables and Measurement Error: Algebra and Statistics. *Journal of Marketing Research*, 18, 382-388. http://dx.doi.org/10.2307/3150980

Gagnon, M. P., Desmartis, M., Labrecque, M., Car, J., Pagliari, C., Pluye, P., Frémont, P., Gagnon, J., Tremblay, N., & Légaré, F. (2012). Systematic review of factors influencing the adoption of information and communication technologies by healthcare professionals. *Journal of medical systems*, *36*(1), 241–277. https://doi.org/10.1007/s10916-010-9473-4

Gagnon, M. P., Ngangue, P., Payne-Gagnon, J., & Desmartis, M. (2016). m-Health adoption by healthcare professionals: a systematic review. *Journal of the American Medical Informatics Association: JAMIA*, *23*(1), 212–220. https://doi.org/10.1093/jamia/ocv052

Garavand, A., Samadbeik, M., Nadri, H., Rahimi, B., & Asadi, H. (2019). Effective Factors in Adoption of Mobile Health Applications between Medical Sciences Students Using the UTAUT Model. *Methods of information in medicine*, *58*(4-05), 131–139. https://doi.org/10.1055/s-0040-1701607

Gefen, D., Straub, D. W., & Boudreau, M.-C. (2000). Structural Equation Modeling and Regression: Guidelines for Research and Practice. *Communications of the Association for Information Systems*, 4, 1-77. https://doi.org/10.17705/1CAIS.00407

Gillan C. (2019). Meeting of the Minds: Considering the Real Impact of Artificial Intelligence. *Journal of medical imaging and radiation sciences*, *50*(4 Suppl 2), S3–S5. https://doi.org/10.1016/j.jmir.2019.11.137

Gong, B., Nugent, J. P., Guest, W., Parker, W., Chang, P. J., Khosa, F., & Nicolaou, S. (2019). Influence of Artificial Intelligence on Canadian Medical Students' Preference for Radiology Specialty: ANational Survey Study. *Academic radiology*, *26*(4), 566–577. https://doi.org/10.1016/j.acra.2018.10.007

Gupta, R. K., & Kumari, R. (2017). National health policy 2017: an overview. *Jk Science*, *19*(3), 135-136.

Hailu, D. T., Melaku, M. S., Abebe, S. A., Walle, A. D., Tilahun, K. N., & Gashu, K. D. (2025). A modified UTAUT model for acceptance to use telemedicine services and its predictors among healthcare professionals at public hospitals in North Shewa Zone of Oromia Regional State, Ethiopia. *Frontiers in digital health*, *7*, 1469365. https://doi.org/10.3389/fdgth.2025.1469365

Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2006). *Multivariate Data Analysis* (6th ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Hair, J., Ringle, C. and Sarstedt, M. (2011) PLS-SEM: Indeed a Silver Bullet. *Journal of Marketing Theory and Practice*, 19, 139-151. https://doi.org/10.2753/MTP1069-6679190202

Hajesmaeel Gohari, S., & Bahaadinbeigy, K. (2021). The most used questionnaires for evaluating telemedicine services. *BMC Medical Informatics and Decision Making*, *21*, 1–11. https://doi.org/10.1186/s12911-021-01407-y

He, J., Baxter, S. L., Xu, J., Xu, J., Zhou, X., & Zhang, K. (2019). The practical implementation of artificial intelligence technologies in medicine. *Nature medicine*, *25*(1), 30–36. https://doi.org/10.1038/s41591-018-0307-0

Hendrix, N., Veenstra, D. L., Cheng, M., Anderson, N. C., & Verguet, S. (2022). Assessing the Economic Value of Clinical Artificial Intelligence: Challenges and Opportunities. *Value in health: the journal of the International Society for Pharmacoeconomics and Outcomes Research*, *25*(3), 331–339. https://doi.org/10.1016/j.jval.2021.08.015

Hengstler, M., Enkel, E., & Duelli, S. (2016). Applied artificial intelligence and trust—The case of autonomous vehicles and medical assistance devices. *Technological Forecasting and Social Change*, *105*, 105–120. https://doi.org/10.1016/j.techfore.2015.12.014

Holden, R. J., & Karsh, B. T. (2010). The technology acceptance model: its past and its future in health care. *Journal of biomedical informatics*, *43*(1), 159–172. https://doi.org/10.1016/j.jbi.2009.07.002

Ifinedo, P. (2012). Technology Acceptance by Health Professionals in Canada: An Analysis with a Modified UTAUT Model. *2012 45th Hawaii International Conference on System Sciences*, 2937-2946. https://doi.org/10.1109/HICSS.2012.556

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology, 2*(4), 230–243. https://doi.org/10.1136/svn-2017-000101

Kalavani, A., Kazerani, M., & Shekofteh, M. (2018). Acceptance of evidence based medicine (EBM) databases by Iranian medical residents using unified theory of acceptance and use of technology (UTAUT). *Health Policy and Technology*, *7*(3), 287–292. https://doi.org/10.1016/j.hlpt.2018.06.005

Kim, H.-W., Chan, H., & Gupta, S. (2007). Value-Based Adoption of Mobile Internet: An Empirical Investigation. *Decision Support Systems*, *43*, 111–126. https://doi.org/10.1016/j.dss.2005.05.009

Kim, S., Lee, K. H., Hwang, H., & Yoo, S. (2016). Analysis of the factors influencing healthcare professionals' adoption of mobile electronic medical record (EMR) using the unified theory of acceptance and use of technology (UTAUT) in a tertiary hospital. *BMC medical informatics and decision making*, *16*, 12. https://doi.org/10.1186/s12911-016-0249-8

Klingberg, A., Sawe, H., Hammar, U., Wallis, L., & Hasselberg, M. (2020). mHealth for burn injury consultations in a low-resource setting: An acceptability study among health care providers. *Telemedicine and e-Health, 26*(7), 833–840. https://doi.org/10.1089/tmj.2019.0004

Ko, M., Wagner, L., & Spetz, J. (2018). Nursing Home Implementation of Health Information Technology: Review of the Literature Finds Inadequate Investment in Preparation, Infrastructure, and Training. *Inquiry : a journal of medical care organization, provision and financing*, *55*, 46958018778902. https://doi.org/10.1177/0046958018778902

Kohnke, A., Cole, M. L., & Bush, R. G. (2014). Incorporating UTAUT predictors for understanding home care patients' and clinicians' acceptance of healthcare telemedicine equipment. *Journal of Technology Management & Innovation, 9*(2), 29–41. https://doi.org/10.4067/S0718-27242014000200003

Laï, M. C., Brian, M., & Mamzer, M. F. (2020). Perceptions of artificial intelligence in healthcare: findings from a qualitative survey study among actors in France. *Journal of translational medicine*, *18*(1), 14. https://doi.org/10.1186/s12967-019-02204-y

Lam, S. Y., Chiang, J., & Parasuraman, A. (2008). The effects of the dimensions of technology readiness on technology acceptance: An empirical analysis. *Journal of Interactive Marketing*, *22*(4), 19-39. https://doi.org/10.1002/dir.20119

Lee, A. T., Ramasamy, R. K., & Subbarao, A. (2025). Understanding Psychosocial Barriers to Healthcare Technology Adoption: A Review of TAM Technology Acceptance Model and Unified Theory of Acceptance and Use of Technology and UTAUT Frameworks. *Healthcare (Basel, Switzerland)*, *13*(3), 250. https://doi.org/10.3390/healthcare13030250

Lee, D., & Yoon, S. N. (2021). Application of Artificial Intelligence-Based Technologies in the Healthcare Industry: Opportunities and Challenges. *International journal of environmental research and public health*, *18*(1), 271. https://doi.org/10.3390/ijerph18010271

Liang, H. and Xue, Y. (2022). Save face or save life: physicians' dilemma in using clinical decision support systems. *Information Systems Research*, *33*(2), 737-758. https://doi.org/10.1287/isre.2021.1082

Lo, W. L. A., Lei, D., Li, L., Huang, D. F., & Tong, K.-F. (2018). The perceived benefits of an artificial intelligence–embedded mobile app implementing evidence-based guidelines for the self-management of chronic neck and back pain: Observational study. *JMIR mHealth and uHealth, 6*(11), e198. https://doi.org/10.2196/mhealth.9825

Lu, L., Cai, R., & Gursoy, D. (2019). Developing and validating a service robot integration willingness scale. *International Journal of Hospitality Management, 80*, 36–51. https://doi.org/10.1016/j.ijhm.2019.01.005

Lu, Q., & Schulz, P. J. (2024). Physician Perspectives on Internet-Informed Patients: Systematic Review. *Journal of medical Internet research*, *26*, e47620. https://doi.org/10.2196/47620

Luxton, D. D. (2019). Should Watson be consulted for a second opinion? *AMA Journal of Ethics, 21*(2), 131–137. https://doi.org/10.1001/amajethics.2019.131

Martin, K. (2019). Designing ethical algorithms. *MIS Quarterly Executive, 18*(2), 129–142. https://doi.org/10.1007/s10551-018-3921-3

Mbelwa, J., Kimaro, H., & Mussa, B. (2019). Acceptability and use of mobile health applications in health information systems: A case of eIDSR and DHIS2 touch mobile applications in Tanzania. In H. M. Rojas, B. R. Rowe, & T. J. Holt (Eds.), *Information and Communication Technologies for Development. Strengthening Southern-Driven Cooperation as a Catalyst for ICT4D* (pp. 603–614). Springer. https://doi.org/10.1007/978-3-030-18400-1_48

Meinert, E., Alturkistani, A., Brindley, D. *et al.* (2018). Weighing benefits and risks in aspects of security, privacy and adoption of technology in a value-based healthcare system. *BMC Med Inform Decis Mak*, 18, 100. https://doi.org/10.1186/s12911-018-0700-0

Mudgal, S.K., Agarwal, R., Chaturvedi, J., Gaur, R. and Ranjan, N. (2022). Real-world application, challenges and implication of Artificial Intelligence in healthcare: an essay. *Pan African Medical Journal*, *43*, 3. https://doi.org/10.11604/pamj.2022.43.3.33384

Murphy, K., Di Ruggiero, E., Upshur, R., Willison, D.J., Malhotra, N., Cai, J.C., Malhotra, N., Lui, V. and Gibson, J. (2021). Artificial intelligence for good health: a scoping review of the ethics literature. *BMC Medical Ethics*, 22 (1), 14. https://doi.org/10.1186/s12910-021-00577-8

Naik, N., Hameed, B.M.Z., Shetty, D.K., Swain, D., Shah, M., Paul, R., Aggarwal, K., Brahim, S., Patil, V., Smriti, K., Shetty, S., Rai, B.P., Chlosta, P. and Somani, B.K. (2022). Legal and ethical consideration in artificial intelligence in healthcare: who takes responsibility?. *Frontiers in Surgery*, 9, 862322. https://doi.org/10.3389/fsurg.2022.862322

Nazar, M., Alam, M.M., Yafi, E. and Su'Ud, M.M. (2021). A systematic review of human-computer interaction and explainable artificial intelligence in healthcare with artificial intelligence techniques. *IEEE Access*, *9*, 153316-153348. https://doi.org/10.1109/ACCESS.2021.3127881

Orruño, E., Gagnon, M. P., Asua, J., & Ben Abdeljelil, A. (2011). Evaluation of teledermatology adoption by health-care professionals using a modified Technology Acceptance Model. *Journal of telemedicine and telecare*, 17(6), 303–307. https://doi.org/10.1258/jtt.2011.101101

Overgoor, G., Chica, M., Rand, W., & Weishampel, A. (2019). Letting the computers take over: Using AI to solve marketing problems. *California Management Review, 61*(4), 156–185. https://doi.org/10.1177/0008125619859318

Patel, B. N., Rosenberg, L., Willcox, G., Baltaxe, D., Lyons, M., Irvin, J., Rajpurkar, P., Amrhein, T., Gupta, R., Halabi, S., Langlotz, C., Lo, E., Mammarappallil, J., Mariano, A. J., Riley, G., Seekins, J., Shen, L., Zucker, E., & Lungren, M. (2019). Human-machine partnership with artificial intelligence for chest radiograph diagnosis. *NPJ digital medicine*, *2*, 111. https://doi.org/10.1038/s41746-019-0189-7

Pedro, A. R., Dias, M. B., Laranjo, L., Cunha, A. S., & Cordeiro, J. V. (2023). Artificial intelligence in medicine: A comprehensive survey of medical doctor's perspectives in Portugal. *PloS one*, *18*(9), e0290613. https://doi.org/10.1371/journal.pone.0290613

Petersson, L., Larsson, I., Nygren, J. M., Nilsen, P., Neher, M., Reed, J. E., Tyskbo, D., & Svedberg, P. (2022). Challenges to implementing artificial intelligence in healthcare: a qualitative interview study with healthcare leaders in Sweden. *BMC health services research*, *22*(1), 850. https://doi.org/10.1186/s12913-022-08215-8

Petkus, H., Hoogewerf, J., & Wyatt, J. C. (2020). What do senior physicians think about AI and clinical decision support systems: Quantitative and qualitative analysis of data from specialty societies. *Clinical Medicine*, *20*(3), 324–328. https://doi.org/10.7861/clinmed.2019-0317

Pinto Dos Santos, D., Giese, D., Brodehl, S., Chon, S. H., Staab, W., Kleinert, R., Maintz, D., & Baeßler, B. (2019). Medical students' attitude towards artificial intelligence: a multicentre survey. *European radiology*, *29*(4), 1640–1646. https://doi.org/10.1007/s00330-018-5601-1

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: a critical review of the literature and recommended remedies. *The Journal of applied psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Reddy, S., Allan, S., Coghlan, S., & Cooper, P. (2020). A governance model for the application of AI in health care. *Journal of the American Medical Informatics Association, JAMIA*, *27*(3), 491–497. https://doi.org/10.1093/jamia/ocz192

Rogers, W.A., Draper, H. and Carter, S.M. (2021). Evaluation of Artificial Intelligence clinical applications: detailed case analyses show value of healthcare ethics approach in identifying patient care issues. *Bioethics*, 35(7), 623-633. https://doi.org/10.1111/bioe.12885

Romero-Brufau, S., Wyatt, K. D., Boyum, P., Mickelson, M., Moore, M., & Cognetta-Rieke, C. (2020). A lesson in implementation: A pre-post study of providers' experience with artificial intelligence-based clinical decision support. *International journal of medical informatics*, *137*, 104072. https://doi.org/10.1016/j.ijmedinf.2019.104072

Rouidi, M., Elouadi, A. E., Hamdoune, A., Choujtani, K., & Chati, A. (2022). TAM-UTAUT and the acceptance of remote healthcare technologies by healthcare professionals: A systematic review. *Informatics in Medicine Unlocked*, *32*, 101008. https://doi.org/10.1016/j.imu.2022.101008

Sarin, S., Sego, T., & Chanvarasuth, N. (2003). Strategic use of bundling for reducing consumers' perceived risk associated with the purchase of new high-tech products. *Journal of Marketing Theory and Practice, 11*(3), 71–83. https://doi.org/10.1080/10696679.2003.11658503

Schepman, A., & Rodway, P. (2022). The General Attitudes towards Artificial Intelligence Scale (GAAIS): Confirmatory Validation and Associations with Personality, Corporate Distrust, and General Trust. *International Journal of Human–Computer Interaction*, *39*(13), 2724–2741. https://doi.org/10.1080/10447318.2022.2085400

Scott, I. A., van der Vegt, A., Lane, P., McPhail, S., & Magrabi, F. (2024). Achieving large-scale clinician adoption of AI-enabled decision support. *BMJ health & care informatics*, *31*(1), e100971. https://doi.org/10.1136/bmjhci-2023-100971

Seele, P., Dierksmeier, C., Hofstetter, R., & Schultz, M. D. (2021). Mapping the ethicality of algorithmic pricing: A review of dynamic and personalized pricing. *Journal of Business Ethics, 170*, 697–719. https://doi.org/10.1007/s10551-019-04371-w

Senbekov, M., Saliev, T., Bukeyeva, Z., Almabayeva, A., Zhanaliyeva, M., Aitenova, N., Toishibekov, Y. and Fakhradiyev, I. (2020). The recent progress and applications of digital technologies in healthcare: a review. *International Journal of Telemedicine and Applications*, 2020, 1-18. https://doi.org/10.1155/2020/8830200

Sendak, M. P., D'Arcy, J., Kashyap, S., Gao, M., Nichols, M., Corey, K., ... & Balu, S. (2020). A path for translation of machine learning products into healthcare delivery. *npj Digital Medicine, 3*, 90. https://doi.org/10.1038/s41746-020-0302-0

Shamszare, H., & Choudhury, A. (2023). Clinicians' Perceptions of Artificial Intelligence: Focus on Workload, Risk, Trust, Clinical Decision Making, and Clinical Integration. *Healthcare (Basel, Switzerland)*, *11*(16), 2308. https://doi.org/10.3390/healthcare11162308

Sharma, R., Yetton, P., & Crawford, J. (2009). Estimating the Effect of Common Method Variance: The Method—Method Pair Technique with an Illustration from TAM Research. *MIS Quarterly*, *33*, 473–490. https://doi.org/10.2307/20650305

Siala, H. and Wang, Y. (2022). Shifting artificial intelligence to be responsible in healthcare: a systematic review. *Social Science and Medicine (1982)*, *296*, 114782. https://doi.org/10.1016/j.socscimed.2022.114782

Stewart, K. A., & Segars, A. H. (2002). An empirical examination of the concern for information privacy instrument. *Information Systems Research,* *13*(1), 36–49. https://doi.org/10.1287/isre.13.1.36.97

Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: Benefits, risks, and strategies for success. *NPJ Digital Medicine, 3*, 17. https://doi.org/10.1038/s41746-020-0221-y

Sykes, T. A., Venkatesh, V., & Gosain, S. (2009). Model of Acceptance with Peer Support: A Social Network Perspective to Understand Employees' System Use. *MIS Quarterly*, *33*(2), 371–393. https://doi.org/10.2307/20650296

Tao, M., Li, X., Alam, F., Yan, Y., & Chau, T. (2025). Unveiling the Impact of AI Technological Anxiety on the Marketers' Intention to Adopt Generative AI. Journal of Global Information Management, 33, 1–22. https://doi.org/10.4018/JGIM.372177

Thekdi, S., Tatar, U., Santos, J., & Chatterjee, S. (2023). Disaster risk and artificial intelligence: A framework to characterize conceptual synergies and future opportunities. *Risk analysis: an official publication of the Society for Risk Analysis*, *43*(8), 1641–1656. https://doi.org/10.1111/risa.14038

Topol, E.J. (2019). High-performance medicine: the convergence of human and artificial intelligence. *Nature Medicine*, *25*(1), 44-56. https://doi.org/10.1038/s41591-018-0300-7

Turja, T., Aaltonen, I., Taipale, S., & Oksanen, A. (2020). Robot acceptance model for care (RAM-care): A principled approach to the intention to use care robots. *Information & Management, 57*(5), Article 103220. https://doi.org/10.1016/j.im.2019.103220

Valente, F., Paredes, S., Henriques, J., Rocha, T., de Carvalho, P. and Morais, J. (2022). Interpretability, personalization and reliability of a machine learning based clinical decision support system. *Data Mining and Knowledge Discovery*, 36 (3), 1140-1173. https://doi.org/10.1007/s10618-022-00821-8

van der Meulen, M. (2019). *Artificial intelligence as a driver of value in value-based health care systems*. Retrieved from: https://www.doktermartijn.nl/static/app/docs/ebook-ai.pdf

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

Vo, V., Chen, G., Aquino, Y. S. J., Carter, S. M., Do, Q. N., & Woode, M. E. (2023). Multi-stakeholder preferences for the use of artificial intelligence in healthcare: A systematic review and thematic analysis. *Social science & medicine (1982)*, *338*, 116357. https://doi.org/10.1016/j.socscimed.2023.116357

Wang, F., & Preininger, A. (2019). AI in Health: State of the Art, Challenges, and Future Directions. *Yearbook of Medical Informatics*, *28*, 016–026. https://doi.org/10.1055/s-0039-1677908

West, S. G., Finch, J. F., & Curran, P. J. (1995). Structural equation models with nonnormal variables: Problems and remedies. In R. H. Hoyle (Ed.), *Structural equation modeling: Concepts, issues, and applications* (pp. 56–75). Sage Publications, Inc.

Whittlestone, J., Nyrup, R., Alexandrova, A., Dihal, K. and Cave, S. (2019) Ethical and Societal Implications of Algorithms, Data, and Artificial Intelligence: A Roadmap for Research. http://www.nuffieldfoundation.org/sites/default/files/files/Ethical-and-Societal-Implications-of-Data-and-AI-report-Nuffield-Foundat.pdf

Wills, M. J., El-Gayar, O. F., & Bennett, D. (2008). Examining healthcare professionals' acceptance of electronic medical records using UTAUT. Issues in Information Systems, 9(2), 396-401. https://doi.org/10.48009/2_iis_2008_396-401

Zheng, R., Jiang, X., Shen, L., He, T., Ji, M., Li, X., & Yu, G. (2025). Investigating Clinicians' Intentions and Influencing Factors for Using an Intelligence-Enabled Diagnostic Clinical Decision Support System in Health Care Systems: Cross-Sectional Survey. *J Med Internet Res*, *27*, e62732. https://doi.org/10.2196/62732

Appendix 2. Constructs Operationalization

Constructs Operationalization on Esmaeilzadeh's framework (2020)

Refernece: **Esmaeilzadeh, P. (2020). Use of AI-based tools for healthcare purposes: a survey study from consumers' perspectives. BMC Medical Informatics and Decision Making, 20(1), 170.**			**Final Rewording**
Construct	**Wording**	**Item Name**	**New Wording**
Perceived performance anxiety (Sarin *et al*., 2003) **(H1)**	I am concerned that the mechanisms used by AI-based devices may lead to inaccurate predictions	PRA1	I am concerned that the mechanisms used by AI-based devices may lead to inaccurate predictions
	I am concerned that the mechanisms used by AI-based devices may result in medical errors	PRA2	I am concerned that the mechanisms used by AI-based devices may result in medical errors
	I am concerned that treatments provided by AI devices may be incomplete	PRA3	I am concerned that treatments provided by AI devices may be incomplete
	I am concerned that the predictive models of AI-based tools may malfunction	PRA4	I am concerned that the predictive models of AI-based tools may malfunction
	I am concerned that the medical decisions made by AI devices may be inadequate	PRA5	I am concerned that the medical decisions made by AI devices may be inadequate
Perceived social biases (Reddy *et al*., 2020) **(H5)**	I am concerned that the AI-based devices may overestimate or underestimate health risks in a certain patient population (e.g., people with insufficient data in AI datasets)	PSB1	I am concerned that the AI-based devices may overestimate or underestimate health risks in a certain patient population (e.g., people with insufficient data in AI datasets)
	I am concerned that data used in the AI devices may lead to societal discrimination to a certain patient group (e.g., minority groups)	PSB2	I am concerned that data used in the AI devices may lead to societal discrimination to a certain patient group (e.g., minority groups)
	I am concerned that AI-based tools used in healthcare may be unfair to a certain group of population (e.g., people with poor access to health care)	PSB3	I am concerned that AI-based tools used in healthcare may be unfair to a certain group of population (e.g., people with poor access to health care)
	I am concerned that AI devices could lead to morally flawed practices in health care	PSB4	I am concerned that AI devices could lead to morally flawed practices in health care
	Overall, I am concerned that the possibility of biases by AI devices to certain groups of the population is high	PSB5	Overall, I am concerned that the possibility of biases by AI devices to certain groups of the population is high
Perceived privacy concerns (Stewart and Segars, 2002) **(H3)**	I think using AI-based applications helps health care entities collect too much personal information from me (original) I think using AI-based applications helps health care entities collect too much personal information **from patients (modified)**	PPC1	I think using AI-based applications helps health care entities collect too much personal information from patients
	I think in this case, I am concerned that health care entities use my health information for other purposes without my knowledge and authorization (original) I think in this case, I am concerned that health care entities use **patients'** health information for other purposes without **their** knowledge and authorization **(modified)**	PPC2	I think in this case, I am concerned that health care entities use patients' health information for other purposes without their knowledge and authorization

	In this case, I am concerned that my health information will be shared with other entities without my explicit consent (original) In this case, I am concerned that **patients'** health information will be shared with other entities without **their** explicit consent **(modified)**	PPC3	In this case, I am concerned that patients' health information will be shared with other entities without their explicit consent
	In this case, I am concerned that unauthorized people will have access to my health information (original) In this case, I am concerned that unauthorized people will have access to **patients'** health information **(modified)**	PPC4	In this case, I am concerned that unauthorized people will have access to patients' health information
	In this case, I am concerned about the privacy of my health information during AI-based health practices (original) In this case, I am concerned about the privacy of **patients'** health information during AI-based health practices **(modified)**	PPC5	In this case, I am concerned about the privacy of patients' health information during AI-based health practices
	In this case, I am concerned my health information would be sold to others without my permission (original) In this case, I am concerned **patients'** health information would be sold to others without **their** permission **(modified)**	PPC6	In this case, I am concerned patients' health information would be sold to others without their permission
Perceived mistrust in AI mechanisms (Luxton, 2019) **(H4)**	I trust in the AI-based clinical tools used for healthcare delivery	PMT1	I trust in the AI-based clinical tools used for healthcare delivery
	I trust in the AI algorithms used in the healthcare	PMT2	I trust in the AI algorithms used in the healthcare
	I trust in AI 's predictive and diagnostic ability for treatment purposes	PMT3	I trust in AI 's predictive and diagnostic ability for treatment purposes
	I trust in the accuracy and predictive powers of current AI algorithmic models used in the medical context	PMT4	I trust in the accuracy and predictive powers of current AI algorithmic models used in the medical context
	I trust that AI-based tools can adapt to specific and unforeseen medical situations	PMT5	I trust that AI-based tools can adapt to specific and unforeseen medical situations
Perceived communication barriers (Lu *et al*., 2019) **(H2)**	I am concerned that AI tools may eliminate the contact between healthcare professionals and patients	PCB1	I am concerned that AI tools may eliminate the contact between healthcare professionals and patients
	I am concerned that AI tools may reduce conversation between physicians and patients	PCB2	I am concerned that AI tools may reduce conversation between physicians and patients
	I am concerned that AI devices may decrease human-aspects of relations in the medical contexts	PCB3	I am concerned that AI devices may decrease human- aspects of relations in the medical contexts
	I am concerned that by using AI devices, I may lose face-to-face cues and personal interactions with physicians (original) I am concerned that by using AI devices, I may lose face-to-face cues and personal interactions with **patients (modified)**	PCB4	I am concerned that by using AI devices, I may lose face-to-face cues and personal interactions with patients

	I am concerned that by using AI devices, I may be in a more passive position for making medical decisions	PCB5	I am concerned that by using AI devices, I may be in a more passive position for making medical decisions
Perceived unregulated standard (Cath, 2018) **(H6)**	I am concerned that special policies and guidelines for AI tools are not transparent yet	PUS1	I am concerned that special policies and guidelines for AI tools are not transparent yet
	I am concerned that the safety and efficacy of AI medical tools are not regulated clearly	PUS2	I am concerned that the safety and efficacy of AI medical tools are not regulated clearly
	I am concerned that regulatory standards to assess AI algorithmic safety are yet to be formalized	PUS3	I am concerned that regulatory standards to assess AI algorithmic safety are yet to be formalized
	I am concerned that appropriate regulatory and accreditation system regarding AI-based devices is not in place yet	PUS4	I am concerned that appropriate regulatory and accreditation system regarding AI-based devices is not in place yet
	I am concerned about the lack of clear guidelines to monitor the performance of AI tools in the medical context	PUS5	I am concerned about the lack of clear guidelines to monitor the performance of AI tools in the medical context
Perceived liability issues (Laï *et al*., 2020) **(H7)**	I am concerned because it is not clear who is responsible when errors result from the use of AI clinical tools	PL1	I am concerned because it is not clear who is responsible when errors result from the use of AI clinical tools
	I am concerned about the liability of using AI-based services for my healthcare (original) I am concerned about the liability of using AI-based services for healthcare **(modified)**	PL2	I am concerned about the liability of using AI-based services for healthcare
	I am concerned because it is not clear who becomes responsible if AI-based tools offer wrong recommendations	PL3	I am concerned because it is not clear who becomes responsible if AI-based tools offer wrong recommendations
	I am concerned because it is unclear where the lines of responsibility begin or end when AI devices guide clinical care	PL4	I am concerned because it is unclear where the lines of responsibility begin or end when AI devices guide clinical care
	I am concerned because it is not clear who is responsible if appropriate AI-recommended treatment options are mistakenly dismissed	PL5	I am concerned because it is not clear who is responsible if appropriate AI-recommended treatment options are mistakenly dismissed
	Overall, I am concerned that the use of AI clinical tools for clinical purposes increases my liability	PL6	Overall, I am concerned that the use of AI clinical tools for clinical purposes increases my liability
Perceived risks (very low/very high) (Bansal *et al*., 2016) **(H8)**	The risk of using AI-based tools for medical purposes is	PR1	The risk of using AI-based tools for medical purposes is
	The degree of uncertainty associated with the use of AI clinical tools is	PR2	The degree of uncertainty associated with the use of AI clinical tools is
	The potential loss associated with the use of AI devices is	PR3	The potential loss associated with the use of AI devices is

	The likelihood of unexpected problems with the use of AI devices is	PR4	The likelihood of unexpected problems with the use of AI devices is
	Overall, the chance of adverse consequences associated with the use of AI-based tools for healthcare purposes is	PR5	Overall, the chance of adverse consequences associated with the use of AI-based tools for healthcare purposes is
Perceived benefits (Lo *et al*., 2018) **(H9)**	I believe AI-based services can improve **diagnostics**	PB1	I believe AI-based services can improve diagnostics
	I think AI-based devices can enhance **prognosis**	PB2	I think AI-based devices can enhance prognosis
	I believe AI-based devices can advance **patient management systems**	PB3	I believe AI-based devices can advance patient management systems
	I believe AI-based tools can suggest accurate **care planning**	PB4	I believe AI-based tools can suggest accurate care planning
	I think AI-based services can recommend reliable **treatment** options	PB5	I think AI-based services can recommend reliable treatment options
	I think AI-based tools can **reduce** healthcare **costs**	PB6	I think AI-based tools can reduce healthcare costs
	Overall, I think AI-based devices can boost **healthcare outcomes**	PB7	Overall, I think AI-based devices can boost healthcare outcomes
Intention to use AI-based tools	I agree to use AI-based tools for clinical purposes	INT1	I agree to use AI-based tools for clinical purposes
	Using AI-based tools for healthcare purposes is something I would consider	INT2	Using AI-based tools for healthcare purposes is something I would consider
	I would like to use AI-based devices to manage my healthcare (original) I would like to use AI-based devices to manage **my patients'** healthcare **(modified)**	INT3	I would like to use AI-based devices to manage my patients' healthcare
	In the future, I am willing to use AI-based services for diagnostics and treatments	INT4	In the future, I am willing to use AI-based services for diagnostics and treatments
	I am very likely to use recommendations provided by AI-based tools for care planning	INT5	I am very likely to use recommendations provided by AI-based tools for care planning
Perception of Workload (PW) Choudhury & Asan (2022) adaptation on NASA-TLX (Task Load Index) by Hart &Staveland (1988)	Using healthcare AI will streamline my clinical workload	PW1	Using healthcare AI will streamline my clinical workload
	I believe that healthcare AI will reduce my overall workload	PW2	I believe that healthcare AI will reduce my overall workload
Perceived AI Risk (R) Choudhury & Asan (2022)	I think using AI for my clinical work will put my patients at risk (reduce patient safety) Using AI for my clinical work will put my patients at a risk which is **(modified)**	R1	Using AI for my clinical work will put my patients at a risk which is
	I think using AI will put my patients' privacy at risk Using AI will put my patients' privacy at a risk which is **(modified)**	R2	Using AI will put my patients' privacy at a risk which is
	I think using AI would improve my clinical decision-making skills/abilities	DM1	I think using AI would improve my clinical decision-making skills/abilities

Perception of Decision Making (DM)	I think using AI would confuse me and hinder my clinical decision-making skills	DM2	I think using AI would confuse me and hinder my clinical decision-making skills

Construct Operationalization on UTAUT-framework (Venkatesh *et al.*, 2003)

Construct	Wording	Source
Facilitating Conditions	I think that my centre has the necessary infrastructure to support my use of teledermatology; I think that my centre has the necessary infrastructure to support my use of **AI (modified)**	Orruño *et al.*, 2011; Gagnon *et al.*; Klingberg *et al.*, 2019; Alabdullah *et al.*, 2020; Mbelwa *et al.*, 2019; Kohnke *et al.*, 2014
	I would use teledermatology if I received adequate training I would use **AI** if I received adequate training **(modified)**	
	I would use teledermatology if I received technical assistance when I needed it I would use **AI** if I received technical assistance when I needed it **(modified)**	
	I have the resources necessary to use such an app I have the resources necessary to use such **a system (modified)**	
	I have the knowledge necessary to use an app like this I have the knowledge necessary to use **a system** like this **(modified)**	
	An app like this is not compatible with the way we work **A system** like this is not compatible with the way we work **(modified)**	
	A specific person (or group) should be available for assistance with difficulties concerning an app like this A specific person (or group) should be available for assistance with difficulties concerning **a system** like this **(modified)**	
	I have resources. (e.g. mobile phone, computer, reporting forms, internet) to use mobile health application I have resources. (e.g. mobile phone, computer, reporting forms, internet) to use an **AI-embedded Clinical Decision Support Systems (modified)**	
	I have the knowledge necessary to use mobile health application I have the knowledge necessary to use an **AI-embedded Clinical Decision Support Systems (modified)**	
	Mobile health application experts are available at any time for assistance with mobile health application difficulties **AI-embedded Clinical Decision Support Systems** experts are available at any time for assistance with AI application difficulties	
	I have knowledge sources (e.g. manuals, documents) to support my use of mobile health application I have knowledge sources (e.g. manuals, documents) to support my use of **AI-embedded Clinical Decision Support Systems (modified)**	
	I have the knowledge necessary to use Telehealth equipment I have the knowledge necessary to use **AI** equipment **(modified)**	
Performance Expectancy	I find mobile health application more useful in my job I find **AI-embedded Clinical Decision Support Systems** more useful in my job **(modified)**	(Jimmy *et al.*, 2019; Jafar *et al.*,

	Using mobile health application increases my productivity Using **AI-embedded Clinical Decision Support Systems** increases my productivity **(modified)**	2020; Kohnke *et al*., 2014)
Effort Expectancy	My interaction with mobile health application is understandable and clear My interaction with **AI-embedded Clinical Decision Support Systems** is understandable and clear **(modified)**	(Jimmy *et al*., 2019; Jafar *et al*., 2020; Kohnke *et al*., 2014)
	Learning on using mobile health application is easy for me Learning on using **AI-embedded Clinical Decision Support Systems** is easy for me **(modified)**	
	It is easy for me to become skillful at using mobile health application It is easy for me to become skillful at using **AI-embedded Clinical Decision Support Systems** (modified)	
	In general, I find mobile health application easy to use In general, I find **AI-embedded Clinical Decision Support Systems** easy to use	
	I expect to become skilled at using the Telehealth equipment I expect to become skilled at using the **AI** equipment **(modified)**	
Social Influence	People who are important to me at work may think that I should use an app like this People who are important to me at work may think that I should use **a system** like this **(modified)**	(Jafar *et al*., 2020; Kohnke *et al*., 2014)
	The senior management of this facility will be helpful in the use of such an app The senior management of this facility will be helpful in the use of such **a system (modified)**	
	In general, the facility management will be supportive of the use of an app of this kind In general, the facility management will be supportive of the use of **a system** of this kind **(modified)**	
	In general, the district health services management will be supportive of the use of such an app In general, the district health services management will be supportive of the use of such **a system (modified)**	
	My coworkers think that I should use mobile health application My coworkers think that I should use **AI-embedded Clinical Decision Support Systems (modified)**	
Training Adequacy	The training on using mobile health application is very helpful in my use of the mobile system The training on using **AI systems** is very helpful in my use of **it (modified)**	(Jimmy *et al*., 2019)
	I feel training received is adequate for my efficient use of mobile health application I feel training received is adequate for my efficient use of **AI-embedded Clinical Decision Support Systems (modified)**	

	I need another training on mobile health application to enable me use the system efficiently I need another training on **AI-embedded Clinical Decision Support Systems** to enable me use the system efficiently **(modified)**	
Behavioural Intention	I have the intention to use telemedicine when it is available at my job I have the intention to use **AI-embedded Clinical Decision Support Systems** when it is available at my job **(modified)**	(Calvo *et al*, 2017)
	I have the intention to use telemedicine when it is necessary to provide health care for my patients I have the intention to use **AI-embedded Clinical Decision Support Systems** when it is necessary to provide health care for my patients **(modified)**	
	I have the intention to use telemedicine regularly with my patients I have the intention to use **AI-embedded Clinical Decision Support Systems** regularly with my patients **(modified)**	

Construct Integration for Longitudinal Analysis (Petkus *et al*., 2020)

Topic Alignment and Gaps Analysis					
Thematic Group	**Theme / Concern**	**Petkus *et al*. (2020)**	**Esmaeilzadeh (2020) Coverage**	**Integration**	**Content**
Performance & Evidence Quality	Accuracy of advice	a) The accuracy of advice may be insufficient for clinical benefit	✓ Already covered: "inaccurate predictions" & "medical errors"	No	
	Clinical effectiveness testing	b) How extensively clinical effectiveness has been tested	✗ Not explicitly phrased this way	Yes	I am concerned that the clinical effectiveness of CDSS has not been sufficiently tested
	Latest evidence base	c) Whether CDSS are based on the latest evidence	✗ Not explicitly phrased this way	Yes	I am concerned that CDSS recommendations may not be based on the latest clinical evidence
Usability & Workflow Integration	Interface/layout consistency	d) The layout of CDSS screens varies between systems, making errors likely	✗ Not covered	Yes	I am concerned that variation in CDSS screen layouts between systems may increase the likelihood of errors
	Data quality dependency	e) CDSS require high-quality, coded data that is rarely available	✗ Not covered	Yes	I am concerned that CDSS require high-quality, coded patient data that is often unavailable
	Output clarity/interpretability	f) CDSS output is not worded clearly and can be hard to interpret	✗ Not covered	Yes	I am concerned that CDSS outputs are not clearly worded and may be hard to interpret
	Speed/workflow disruption	g) CDSS can be too slow, disrupting workflows	✗ Not covered	Yes	I am concerned that CDSS tools may be too slow and disrupt clinical workflows
	Patient preferences consideration	h) CDSS can ignore patient preferences	✗ Not covered	Yes	I am concerned that CDSS may ignore patient preferences.
Performance & Evidence Quality	Validation obsolescence	i) Validation studies can become obsolete as algorithms evolve	✗ Not phrased exactly this way	Yes	I am concerned that CDSS validation studies may quickly become obsolete as algorithms evolve.
	Validation depends on data quality	j) Validation study outcomes depend on data quality	✗ Not covered	Yes	I am concerned that the reliability of CDSS validation depends heavily on the quality of patient data.
	Scope of use clarity	k) Some CDSS don't describe their intended aim/scope	✗ Not covered	Yes	I am concerned that CDSS do not clearly describe their intended scope of use.
Training & Clinical Judgment	User competence clarity	l) Some CDSS don't describe the level of user knowledge/experience required	✗ Not covered	Yes	I am concerned that CDSS do not specify the level of user knowledge or experience required for safe use.

Regulatory Standards	Regulatory rigor	m) Regulatory rigor for CDSS is too lax	✓ Covered via regulatory concerns	No	
	Regulatory transparency (peer-reviewed testing)	n) Testing standards for CDSS are not peer-reviewed	✗ Not covered	Yes	I am concerned that CDSS testing standards are not open or peer reviewed.
	Open publication of testing results	o) Testing results for CDSS should be openly available	✗ Not covered	Yes	I am concerned that the results of CDSS testing are not openly published.
Cross-Cutting Ethical Concerns	Unconscious bias	p) CDSS can embed unconscious bias	✓ Covered via social biases	No	
	Black-box opacity	q) CDSS as “black boxes” lacking user insight	✓ Covered: "black box problem"	No	
Training & Clinical Judgment	Legal liability unclear	r) Legal liability of doctors unclear	✓ Covered	No	
Usability & Workflow Integration	Alert fatigue	s) CDSS produce too many alerts/reminders	✗ Not covered	Yes	I am concerned that CDSS tools may generate excessive alerts, leading to alert fatigue.
Training & Clinical Judgment	Over-reliance/misdirection	t) Doctors may follow incorrect CDSS advice	✗ Not covered	Yes	I am concerned that doctors may sometimes follow incorrect CDSS advice, even if they would otherwise make correct decisions.
	Juniors follow uncritically	u) CDSS may lead juniors to follow guidance uncritically	✗ Not covered	Yes	I am concerned that CDSS may lead junior staff to follow guidance uncritically.
	Reduced training opportunities for juniors	v) CDSS reduces training for juniors	✗ Not covered	Yes	I am concerned that the use of CDSS may reduce training opportunities for junior clinicians.
Regulatory Standards	Scarce expert input in system design/updates	w) Scarce expert input during CDSS development	✗ Not covered	Yes	I am concerned that there is insufficient expert clinical input in the design, implementation, and updating of CDSS systems.
Thematic Group	**Theme / Benefit**	**Petkus *et al*. (2020)**	**Esmaeilzadeh (2020) Coverage**	**Integration**	**Content**
Patient-Centered Outcomes	Improved patient safety	a) Improved patient safety	✗ Not explicitly included	Yes	I believe AI-based CDSS can improve patient safety
Clinical Efficiency & Resource Use	More efficient clinical work	b) More efficient clinical work	✓ UTAUT: "increases productivity", "accomplish tasks quickly"	No	

	Reduced resource utilization	c) Reduced resource utilization	✓ Esmaeilzadeh: "reduce healthcare costs"	No	
Patient-Centered Outcomes	Better patient outcomes	d) Better patient outcomes	✓ Esmaeilzadeh: "boost healthcare outcomes"	No	
Clinical Efficiency & Resource Use	Improved medicines management	e) Improved medicines management	✗ Not explicitly included	Yes	I believe AI-based CDSS can improve medicines management (e.g., dosage accuracy, prescription safety)
	Better use of investigations	f) Better use of investigations	✗ Not explicitly included	Yes	I believe AI-based CDSS can facilitate better use of clinical investigations (e.g., more appropriate ordering of tests)
Patient-Centered Outcomes	More accurate diagnoses	g) More accurate diagnoses	✓ Esmaeilzadeh: "improve diagnostics"	No	
	Fewer drug side effects	h) Fewer drug side effects	✗ Not explicitly included	Yes	I believe AI-based CDSS can reduce drug side effects by supporting appropriate prescribing
	More attention to preventive care	i) More attention to preventive care	✗ Not explicitly included	Yes	I believe AI-based CDSS encourages more attention to preventive care
System Readiness & User Support	Training adequacy	a) UTAUT/Training block	✓ Included	No	
	Facilitating conditions (resources, support)	b) UTAUT block	✓ Included	No	
	Performance expectancy (usefulness, productivity)	c) UTAUT survey	✓ Included	No	

Chapter 5. **NHS Procurement Preferences for Trustworthy AI: A Discrete Choice Experiment**

Simona Curiello[1, *], Adriane Chapman[2], Jeremy Wyatt[3]

[1] *Department of Social Sciences, University of Foggia, Italy*

[2] *Digital Health and Biomedical Engineering, University of Southampton, UK*

[3] *School of Primary Care, Population Sciences and Medical Education, University of Southampton, UK*

[*] Corresponding Author. Email: simona.curiello@unifg.it

NHS Procurement Preferences for Trustworthy AI: A Discrete Choice Experiment

Abstract

Despite the increasing availability of regulatory-approved artificial intelligence (AI) tools, their adoption within the UK National Health Service (NHS) remains low. This study addresses a critical gap in understanding how NHS professionals make procurement decisions for AI-enabled Clinical Decision Support Systems (AI-CDSSs), a process that is frequently disregarded in favour of studying design, implementation, or post-deployment evaluation. The research focuses on procurement within the UK National Health Service (NHS), where purchasing decisions are shaped by a complex balance between fiscal accountability, ethical oversight, and regulatory compliance. Using a mixed-method approach that combines think-aloud protocols with Choice-Based Conjoint (CBC) analysis, we investigate how NHS-affiliated professionals evaluate and trade off key AI-CDSSs attributes. A systematic literature review identified 44 attributes, conceptualized as key features or characteristics that influence decision-makers' preferences. Through a series of expert consultations with relevant stakeholders, this number was methodically pruned to result in a final selection of five attributes: system usability, algorithmic fairness, diagnostic accuracy, explainability, and ease of adaptability to NHS standards. 33 NHS professionals were given 12 choice tasks replicating actual procurement scenarios as part of the discrete choice experiment (DCE). Early results showed that the most important factors were algorithmic fairness, explainability, and accuracy. In contrast, usability and ease of adaptability to standards exert a significantly lesser influence. These initial findings offer important insights for guiding vendor strategies, informing NHS procurement policies and enhancing institutional readiness for trustworthy AI integration.

Introduction

The AI field is currently experiencing a period of rapid growth and development, with CDSSs representing one of the most promising applications. The utilisation of AI-enabled tools has the potential to enhance diagnostic accuracy and patient outcomes, optimise resource allocation, and streamline clinical workflows. However, despite the growing body of academic literature promoting the benefits of AI in healthcare, its actual implementation remains modest across many clinical settings – particularly within publicly funded health systems such as the UK National Health Service (NHS) (Zhang *et al*., 2022; Muehlematter *et al*., 2021). While more than 950 AI or machine learning (ML)-enabled medical devices have been authorized by the U.S. Food and Drug Administration (FDA) and over 240 CE-marked devices were approved in Europe between 2015 and 2020 (Badnjevic *et al*., 2021; Muehlematter *et al*., 2021), these figures have not translated into the same proportionate adoptions within NHS hospitals and trusts.

The NHS is one of the largest single-payer health services in the world, and its procurement ecosystem is strategically positioned to determine whether and how new technologies reach clinical practice (Grammatopoulos *et al*., 2024). NHS procurement is regulated through a mixture of value-based purchasing, ethical oversight, and regulatory compliance, reflecting both fiscal constraints and the need to safeguard patient outcomes (Meehan *et al*., 2017). However, the procurement of AI technologies is fundamentally different from that of traditional medical devices, because of the involvement of algorithmic systems whose performance can evolve over time and which raise issues of explainability, accountability, and data governance (Daye *et al*., 2022). Recent reviews have highlighted the need for NHS procurement teams to consider not only technical performance and cost but also issues such as algorithmic fairness, interoperability with existing EHR infrastructures, and compliance with frameworks such as the NHS AI Lab's "*Buyer's Guide to AI in Health and Care*" (NHSX, 2021). To prove compliance with safety and performance standards set by the Medicines and Healthcare Products Regulatory Agency, all AI-enabled medical devices must receive the UK Conformity Assessed (UKCA) marking (MHRA, 2023). The Evidence Standards Framework (ESF) from the National Institute for Health and Care Excellence also helps developers and purchasers assess the practical performance, economic worth, and efficacy of digital health technologies (NICE, 2022). Together, these frameworks seek to operationalize the value and risks of adopting AI in NHS settings in a trustworthy manner, bringing regulatory compliance into line with patient-centred and evidence-based decision-making.

A fundamental explanation for the development-deployment discrepancy lies in the fragmented understanding of the implementation pipeline, especially from the perspective of healthcare procurement professionals (Goldfarb & Teodoridis, 2022). Existing studies have tended to concentrate on the design of algorithms, clinical utility, or post-deployment monitoring, while neglecting procurement – a critical, yet often overlooked, phase that determines whether and how AI-CDSSs tools are adopted within healthcare institutions (Gama *et al*., 2022). The decision-making process involved in procurement is inherently complex, necessitating trade-offs across a range of factors including technical specifications, ethical standards, cost-effectiveness, regulatory compliance, and long-term integration risks (Kim *et al*., 2023). In practice, NHS procurement professionals work at the intersection of clinical priorities, regulatory scrutiny, and budgetary accountability, making their preferences essential to comprehend how AI innovation is translated into operational reality ((Office for Artificial Intelligence *et al*., 2020). According to recent studies of NHS digital procurement, the pathway from innovation to implementation is often hampered by a "last-mile problem", when promising AI solutions fail to overcome pilot stages because of hazy value

propositions, risk aversion, or a lack of alignment with institutional procedures (Chada & Summers, 2022; Evans *et al*., 2025).

Despite the centrality of procurement to AI adoption, empirical research rarely focuses on the pivotal role played by procurement professionals in the realm of healthcare innovation, nor on their perceptions of the attributes (characteristics) that most influence their preferences during the procurement process of a new CDSS (Petersson *et al*., 2022).

To address this gap, the present study adopts a mixed-method design that combines the quantitative rigor of Choice-Based Conjoint (CBC) analysis with qualitative insights derived from think-aloud protocols and semi-structured interviews. Specifically, the research seeks to answer the following question:

What are the key factors and trade-offs influencing procurement decisions for AI-CDSSs in the NHS, and how do front line NHS staff perceive these trade-offs in practice?

The choice-based experiment facilitates a systematic elicitation of procurement preferences across pre-defined attributes, while the qualitative component provides contextual depth by capturing how decision-makers interpret trade-offs, weigh organisational priorities, and navigate uncertainty (Merriam & Tisdell, 2015).

This approach is particularly appropriate due to the limited research on AI adoption from the procurement perspective. Moreover, it responds to recent recommendations advocating for a structured, end-to-end understanding of AI implementation that includes procurement as a critical phase and connects it with prior evaluation and post-deployment monitoring activities (de Hond *et al*., 2022; Wu *et al*., 2021).

Theoretical Background

Studies suggest that individuals formulate their preferences dynamically, influenced by the context in which they are making decisions, rather than holding pre-existing fixed preferences (Louviere *et al*., 2007; Van Houtven *et al*., 2011; Yeh *et al*., 2020). Behavioural economics theories also indicate that when facing real-time decision-making scenarios, people often display behaviours such as anxiety and inconsistency over time (Roosen & Marette, 2011). Additionally, decision-makers (DMs) might shy away from complex choices due to self-control limitations, leading to cognitive biases, especially under uncertain conditions (Kahneman & Tversky, 1979).

In the realm of consumer behaviour research, choice-based methods are frequently used to determine which product attributes are most significant (Ben-Akiva *et al*., 2019). However, existing literature in clinical decision-making often overlooks the potential of choice-based analyses like

Conjoint Analysis (CA) to effectively capture decision-makers' preferences in healthcare settings (Mele, 2008). CA is a well-established method in marketing research for examining consumer preferences through survey-based data (Mühlbacher & Johnson, 2016). This approach applies statistical techniques to evaluate preference responses, estimating preference weights (part-worth utilities), and employs choice simulators – software tools that simulate how people might choose between different product or service configurations – to predict how individuals make selections from various option sets across a population (Mele, 2008; Cunningham *et al*., 2010; Louder *et al*., 2016).

In most decision-making situations, people do not typically evaluate or rank all available options from best to worst; rather, they tend to select one option directly. Choice-Based Conjoint (CBC) analysis, a variation of CA, is a survey-based method designed to uncover how individuals navigate complex decisions that involve balancing trade-offs between competing alternatives (Raghavarao *et al*., 2010).

CBC has been widely adopted in various fields, including marketing, healthcare, transportation, and information systems research (Rao, 2014; Sharma *et al*., 2023). Within CBC studies, participants are shown multiple sets of alternatives (known as *choice sets*) and asked to select the alternative that offers the greatest utility in each set. These choice sets are crafted using a fractional factorial orthogonal experimental design, which creates hypothetical decision-making scenarios by combining different product or service attributes in structured ways for evaluation (Wyatt *et al*., 2010).

For example, when determining the best Clinical Decision Support (CDS) tool to support patient managment, healthcare professionals often select the option they perceive as most accurate and clinically trustworthy (Kawamoto *et al*., 2005; Dellink, 2023). The decision process itself may be multi-faceted, incorporating both cognitive reasoning and evidence-based considerations. In such scenarios, CBC methodology is particularly useful in identifying the preferred choice among a set of potential options, aiming to maximize the decision-maker's utility or satisfaction.

Compared to traditional methods like regression analysis (Karakaya & Awasthi, 2015) or the Analytic Hierarchy Process (AHP) (Helm *et al*., 2008), CBC is considered to offer greater robustness in understanding decision-making. While AHP has been identified as one of the most commonly used multi-criteria decision analysis (MCDA) techniques in medical decision-making (Adunlin *et al*., 2015), CBC stands out as a superior tool for capturing consumer preferences when they must choose between multiple competing alternatives (Afful-Dadzie & Egala, 2022).

CBC helps quantify the influence of various attributes on consumer choices, providing preference weights for each factor (Roßnagel *et al*., 2014). Discrete choice models like CBC explicitly account for the attribute levels and ranges, which enables a clearer depiction of the relative importance assigned to different attributes – an area where AHP falls short (Danner *et al*., 2017).

Unlike direct survey methods that rely on self-reporting, CBC is less susceptible to social desirability bias, as it infers preferences from actual trade-offs made in hypothetical but realistic choice scenarios (Huls *et al*., 2023). In contrast to other ranking methods, such as Best-Worst Scaling (BWS) or AHP (Benlian, 2011; Lansing *et al*., 2019), CBC is particularly suited for complex decision-making scenarios where individuals assess multiple attributes and make trade-offs, such as selecting between products or services (Richard *et al*., 2012).

AI-CDSS Procurement in Healthcare

A complex interaction of organizational, regulatory, ethical, and technical variables influences healthcare procurement decisions for the implementation of AI technologies. The importance of the procurement viewpoint in understanding the barriers to AI adoption in the field of medical diagnostics has been emphasized in a work published by Dellink in 2023. The author highlights the necessity to match physician needs and accessible technology with purchase decisions (Dellink, 2023). In a similar vein, Uravirta's (2023) research pinpoints the critical elements impacting the acquisition of machine learning software in the medical field. These include regulatory compliance, vendor reputation, interoperability, and the level of clinician involvement in procurement decisions (Uravirta, 2023). In their 2023 publication, Stogiannos *et al*. focus on AI governance frameworks in the context of UK medical imaging and radiotherapy. The authors conclude that robust governance policies are necessary to guide the procurement and adoption processes and stress that openness, data governance, and clinicians' engagement are all critical elements. The adoption and implementation of AI-powered CDSS remain challenging, as evidenced by numerous reviews that have pointed out several barriers and facilitators. Shakibaei Bonakdeh and Sohal (2024) conducted a meta-synthesis of qualitative studies and identified procurement-related challenges such as unclear return on investment (ROI), integration problems, and clinician scepticism. The prevailing opinion is that to overcome these challenges, vendors and healthcare institutions need to engage in closer collaboration (Shakibaei Bonakdeh & Sohal, 2024). In their study, Finkelstein *et al*. (2024) investigated the use of AI-CDSS within Electronic Health Records (EHRs), and found that facing technical integration challenges, workflow issues, and clinician trust represented key factors for successful adoption (Finkelstein *et al*., 2024). Furthermore, to address the issue of low adoption rates, Bertl *et al*. (2023) suggest a framework of success factors for implementing AI-CDSS. This includes strong leadership, interdisciplinary collaboration, and ongoing evaluation of system performance. Weinert *et al*. (2022) analysed the perspectives of IT decision-makers in several German hospitals and identified organisational readiness, managing logistics, and having centralized purchasing processes as crucial factors influencing AI adoption. Additionally, the findings highlighted the pivotal role of effective change management and training programs in facilitating technology adoption. Several studies have

also highlighted ethical challenges linked to the procurement of AI-CDSSs. Adler-Milstein *et al*. (2022) discuss the tension between procurement choices and clinician trust, advocating for transparency and bias mitigation strategies in AI systems. Finally, Yang *et al*. (2022) used the Technology-Organization-Environment (TOE) framework to investigate the adoption of AI, identifying external pressure, technological complexity, and organizational culture as decisive factors.

These results align with data from real-world assessments of AI-CDSS procurement, where decision-makers and clinicians have identified a set of recurring concerns (Petkus *et al*., 2020). For example, in real-world contexts, clinicians assessing AI tools often prioritize accuracy and usability (von Wedel & Hagist, 2022; Egala & Liang, 2024). Furthermore, algorithmic fairness has become a crucial procurement criterion, especially when it comes to clinical validation and the fair performance of AI systems across diverse patient populations (Latt *et al*., 2024; Ploug *et al*., 2021). Particularly in safety-critical healthcare settings, regulatory compliance – which includes aspects such as certification and auditability – is also considered a fundamental requirement (Woode *et al*., 3025). Finally, studies evaluating digital health technologies using DCEs have identified other factors that influence procurement decisions, including data privacy and update frequency (Zhang *et al*., 2025).

Table 12 – Procurement and Implementation of AI-CDSSs in Healthcare

Theme	Key Topics	Authors
Procurement Determinants	Cost, integration, regulatory standards, vendor reputation	Dellink (2023); Uravirta (2023); Stogiannos *et al*. (2023)
Procurement Challenges	ROI uncertainty, lack of alignment between procurement and clinical needs	Shakibaei Bonakdeh & Sohal (2024)
Implementation Barriers	Workflow integration, clinician trust, technical limitations	Finkelstein *et al*. (2024); Wang *et al*. (2021); Giebel *et al*. (2025)
Ethical & Trust Issues	Algorithm transparency, liability, fairness, trustworthiness	Jones *et al*. (2023); Adler-Milstein *et al*. (2022); Dingel *et al*. (2024)
Facilitators of Adoption	Strong leadership, interdisciplinary collaboration, training and support	Bertl *et al*. (2023); Kočo *et al*. (2024); Schouten *et al*. (2020)
Governance and Frameworks	AI policy, governance models, stakeholder engagement	Stogiannos *et al*. (2023); Bai *et al*. (2025); Yang *et al*. (2022)
Clinician Perception	Perceived utility, explainability, cognitive load, attitudes toward AI in decision-making	Peek *et al*. (2024); Fujimori *et al*. (2022); Burgess *et al*. (2023)

Source: Authors' own elaboration

As noted in HSSIB's thematic review of Electronic Patient Records (EPR) systems (March 2025), some procurement decisions were influenced by strategic or financial pressures rather than actual capability, leading to interoperability failures and costly retrofitting. Such concerns equally apply to AI-CDSS procurement and highlight a growing policy gap: developers frequently promote AI tools to NHS stakeholders without a robust understanding of what procurement professionals value. Decision-makers are often left to trade off highly accurate but opaque systems against less accurate

but more explainable ones – without clear national criteria to support these judgments (Morandini, 2023; Kovári, 2024).

Attributes and Levels Development for Conjoint Analysis

Using the Scopus database, a systematic literature review (SLR) was carried out to determine important characteristics for the DCE. The search focused on technical, organizational, and ethical aspects of the adoption and acquisition of AI-enabled CDSSs in healthcare through peer-reviewed and grey literature. After PRISMA-based screening (Tricco *et al*., 2018), 16 high-quality sources were retained from an initial 502 records. The Consolidated Framework for Implementation Research (CFIR – Coast *et al*., 2012; Bridges *et al*., 2011; de Bekker-Grob *et al*., 2012) served as a guide for attribute selection, yielding 44 potential attributes and five overarching domains: 1) system characteristics & technical design; 2) data infrastructure & quality; 3) user-centred factors (perceptions, experience, capability); 4) implementation process & support and 5) external context & policy environment (see Table 13 for details).

Table 13 – Attributes derived from Systematic Literature Review

1. System Characteristics & Technical Design *Attributes related to the technical features and operability of the AI-CDSS*		**References**
System Usability	Ease of use, minimal disruption to workflows	Syed Nasirin *et al*., 2024
Interface Design	Intuitive interface and navigation	Nair *et al*., 2023
Workflow Integration	Seamless embedding into existing clinical workflows	Syed Nasirin *et al*., 2024
Integration / Interoperability	Compatibility with EHR or HIS, real-time updates	Nair *et al*., 2023; Finkelstein *et al*., 2024
Update Frequency	Adaptation to new guidelines	Nair *et al*., 2023
Alert Fatigue	Interface design to prevent overload	Yoo *et al*., 2023
Customization	Adaptability to regional health and hospital data	Syed Nasirin *et al*., 2024
Automation Safety	Reduction of overreliance	Magrabi *et al*., 2019
Override Functions	User control over AI	Van Biesen *et al*., 2022
Scalability	Tailored local use	Finkelstein *et al*., 2024
Bias Mitigation	Fairness in decision outputs	Janssen *et al*., 2024
Actionable Insights	Understandable risk scores	Nair *et al*., 2023
Clinical Task Fit	Alignment with task complexity	Tokgöz *et al*., 2024
Monitoring	Capability and sufficiency of parameters	Tokgöz *et al*., 2024
2. Data Infrastructure & Quality *Attributes pertaining to the quality, availability, and structure of data required for AI-CDSS functioning*		**References**
Data Compatibility/Data Security	Infrastructure integration/GDPR, privacy standards	Bergquist *et al*., 2024; Parchmann *et al*., 2024
Data Quality	Structured, high-quality inputs for accurate decisions	Syed Nasirin *et al*., 2024
Clinical Performance	False positives, empirical validation	Finkelstein *et al*., 2024
Evidence Base	Empirical validation	Magrabi *et al*., 2019
3. User-Centered Factors (Perceptions, Experience, Capability) *Factors affecting how users perceive, trust, and interact with the system*		**References**
Performance Expectancy	Expected improvements in job efficiency	Ratta *et al*., 2025

Effort Expectancy	Ease of use and minimal training	Ratta *et al*., 2025
Perceived Usefulness	Security certification, data protection	Mosch et al., 2022
Perceived Risk	Access to training resources	Ratta *et al*., 2025
Trust	Trust level influenced by reliability, safety	Ratta *et al*., 2025
Self-Efficacy	Concerns about errors and privacy	Ratta *et al*., 2025
Human-AI Interaction	Avoiding skill erosion	Tokgöz *et al*., 2024
User Acceptance	Ongoing institutional support, including helpdesks	Haque et al., 2023
Professional Identity	Explainability of decision logic	Ratta *et al*., 2025; Bergquist *et al*., 2024
Change Readiness	Training and cultural adaptation	Tokgöz *et al*., 2024
User Readiness	Continuous training, digital literacy	Syed Nasirin *et al*., 2024
Ethical Acceptability	Post-sale tech assistance	Parchmann *et al*., 2024

4. Implementation Process & Support *Operational and support mechanisms enabling rollout, sustainability, and refinement*		**References**
Iterative Evaluation	Continuous refinement via systems engineering	Syed Nasirin *et al*., 2024
Feedback Mechanisms	Pilot phase engagement	Finkelstein *et al*., 2024
Economic Efficiency	Cost-effectiveness, budget constraints, ROI	Syed Nasirin *et al*., 2024
Technical Support	Policy guidance in procurement	Syed Nasirin *et al*., 2024
Vendor Support	Post-sale tech assistance	Yoo *et al*., 2023
Implementation Climate	Liability in override scenarios	Syed Nasirin *et al*., 2024
Socio-Technical Integration	Alignment with EHR	Finkelstein *et al*., 2024
Implementation Frameworks	Real-world adaptation models	Mi *et al*., 2021
Governance	Vendor transparency, policy	Nair *et al*., 2023
Stakeholder Engagement	Involvement of clinicians, IT, and administrators	Syed Nasirin *et al*., 2024

5. External Context & Policy Environment *Institutional, regulatory, and policy elements influencing procurement and governance*		**References**
Transparency & Safety	Trust from transparency and safety	Syed Nasirin *et al*., 2024
Policy Alignment	Policy guidance in procurement	Syed Nasirin *et al*., 2024
Regulatory Compliance	Trust influenced by reliability, safety	Finkelstein *et al*., 2024
Legal Clarity	Alignment with clinical or ICU context	Nair *et al*., 2023

Source: our elaboration

Methodology

According to the existing methodological literature, the selection of attributes for DCEs in healthcare often follows a formalized two-stage process. As explained by Coast *et al*. (2012), the first stage is conceptual attribute generation by exploratory qualitative methods, such as in-depth interviews or semi-structured interviews with relevant stakeholders. This step ensures that the attributes capture real-world problems and decisional scenarios. The second stage involves refining and rewording attributes into well formed, respondent-friendly phrases and defining possible levels for each attribute. The two stages are not always differentiated in some studies, but the process always includes both conceptual and practical refinement. Alternative approaches such as meta-ethnography,

which synthesises qualitative literature on a specific topic, and focus groups have also been used to offer a rigorous and context-sensitive list of DCE attributes (Noblit & Hare, 1988).

To enhance the validity and applicability of the research, an iterative design process was employed, including stakeholder input, industry-specific language, and empirical pre-testing (see *Appendix 3*). Ethical approval was granted by the Faculty Ethics Committee (FEC) of the University of Southampton [ERGO No.: 105265].

Based on their prominence in recent empirical studies and compatibility with the UK healthcare procurement context, the long list of attributes obtained from the SLR was then whittled down to a shortlist of 12 attributes.

To further validate the conceptual and contextual relevance of these 12 refined attributes, members of the Chief Clinical Information Officer (CCIO) Advisory Board were involved in an expert elicitation process. Three main steps were taken in the consultation:

- **Task 1**: Review and prioritize the 12 shortlisted attributes, based on expert judgement about their relevance to AI-CDSS procurement in the NHS.
- **Task 2**: Select a maximum of 5 attributes to include in the DCE to ensure clarity, feasibility, and statistical robustness.
- **Task 3** *(Optional):* Provide feedback to help refine the clarity and contextual alignment with NHS procurement language by rewording and interpreting the chosen attributes and levels.

After that, each attribute was divided into levels that corresponded to conceivable differences as perceived by procurement stakeholders and medical professionals. The final set of attributes, along with definitions and supporting literature, are shown in the Table 14. These serve as the foundation for the DCE survey design:

Table 14 - DCE attributes and attribute levels for AI-based assistance tools

Attribute Name	Description	1st Level	2nd Level	3rd Level
System Usability/ Workflow Integration	Degree to which the AI-CDSS is intuitive and supports clinical tasks with minimal disruption	≤2 clicks, integrates with clinical pathways	3–5 clicks, occasional override	>5 clicks, frequent overrides/disruption to tasks
Algorithmic Fairness/ Bias Mitigation	Extent to which the AI-CDSS performs equitably across diverse patient groups	Validated across multiple demographic groups	Tested in some populations, but limited generalizability	Unknown or unreported bias levels
Accuracy	Credibility and performance of the system's output, based on source and quality of validation	≥95% accuracy, validated in peer-reviewed studies	90–94% accuracy/ without peer-reviewed evidence	<90% accuracy or based only on internal developer validation
Explainability	Transparency of the AI-CDSS decision-making process in ways that clinicians can interpret, trust, and defend professionally	Highly explainable with interpretable models and transparent logic	Moderate explainability; partial transparency or simplified summaries	Opaque system with no understandable rationale for outputs

Ease of Adaptability to Standards	Ability of the AI-CDSS to comply with and adapt to evolving NHS standards	Fully certified + adaptable to NHS/AI Act evolution	Meets current standards (UKCA marked; MHRA standard; NICE ESF) but inflexible	Legacy certifications or unclear liability in case of error

Source: Authors' own elaboration

Subsequently, prior to the initiation of a complete DCE, a pilot group of 3 advisers on AI purchasing were asked to conduct a think-aloud protocol. The semi-structured nature of these sessions was conducive to the fulfilment of several key functions: refining attributes' levels; ensuring alignment with real-world procurement criteria; exploring attributes' relevance; verifying understandability and realism within the realm of stakeholder expectations; using authentic terminology; improving interpretability and face validity of survey items.

Participants were instructed to do the DCE via screen-share (either in person or via Zoom) with a controlled think-aloud protocol. As participants compared hypothetical scenarios, they were prompted to consider that they were invited to provide an explicit explanation of the reasoning process for comparing the two alternatives. This enabled the research team to detect instances of confusion, cognitive shortcuts, or decision fatigue in real time. The think-aloud protocol, as seen by Wright *et al*. (2022) and Queally *et al*. (2021), was discovered to increase DCE task interpretability and inform necessary design refinements.

In conclusion, the preliminary investigation was conducted by a group of content experts who evaluated the instrument's comprehensibility and usability. The multi-stage, user-driven process made the final attributes and levels used in the DCE more conceptually valid and contextually relevant.

Experimental Design for Discrete Choice Modelling

As anticipated, this research utilizes Choice-Based Conjoint (CBC) analysis to explore participants' preferences towards the adoption of CDSSs applications. NHS advisors on AI purchasing were presented with a series of choice scenarios in which they were asked to evaluate and select between procurement offers for AI-based CDSSs from various vendors.

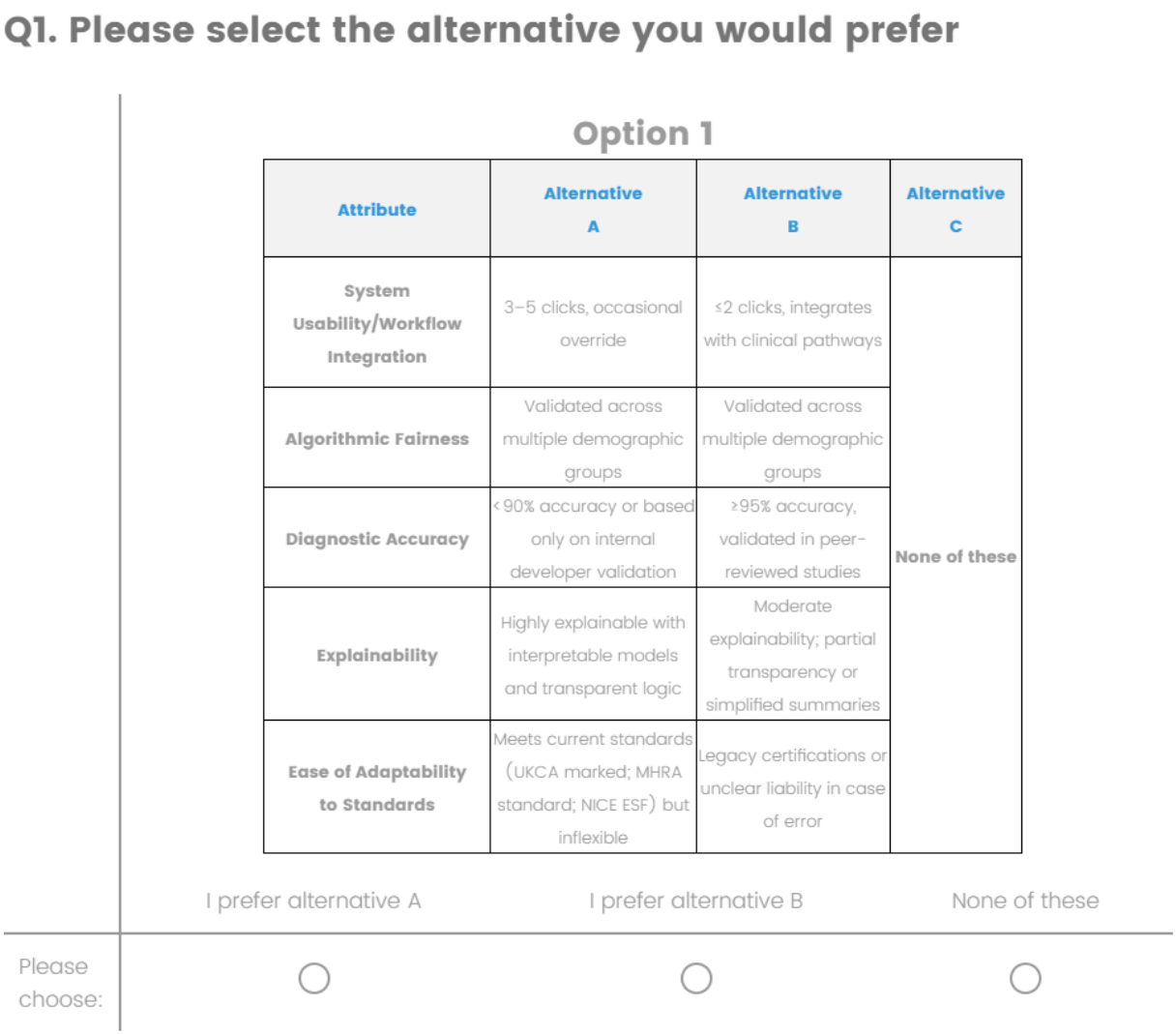
For the present study, a D-efficient experimental design for DCE was developed using JMP 18 (SAS Institute Inc.), thereby ensuring high statistical precision in estimating attribute effects. As anticipated in the paragraph on *Attributes and Levels Development*, five attributes, each with three categorical levels, were entered as factors.

Given that five attributes with three levels each yield a total of $3^5 = 243$ possible profiles and consequently $(243\times 242)/2 = 29403$ potential pairwise combinations, presenting the full factorial design to participants would be impractical. To address this, we implemented a fractional factorial design (Hahn & Shapiro, 1966) that strategically selects the most informative combinations, thereby reducing complexity while preserving statistical robustness. This configuration of experimental setup

is frequently employed to diminish intercorrelations among the various attributes under investigation, thereby ensuring that the influence of each attribute can be estimated independently. Furthermore, this approach serves to minimise overlap or confounding between the effects of various factors, thereby facilitating a more straightforward interpretation of results. A main-effects-only model was selected, in line with standard practice for initial preference elicitation, excluding interaction terms to reduce respondent burden and cognitive load. The fractional factorial design was constructed by specifying 24 total profiles, arranged into 12 choice sets with 2 alternatives per set (excluding the "opt-out" option). This setup allows estimation of main attribute effects while maintaining efficiency and feasibility for respondents.

In accordance with the prevailing best practices in Discrete Choice Experiment (DCE) methodology, a "no-choice" option was incorporated to reflect the reality that decision-makers may elect to forgo procurement entirely if no option meets their standards (Haaijer *et al*., 2001). Each participant was asked to complete 12 choice sets, with each set presenting two CDSSs product configurations and the no-choice alternative (see example in Figure 1). A range of demographic and contextual data were also collected for the purpose of in-depth analysis, including age, gender, professional experience, role in AI purchasing and familiarity with AI.

Figure 1 – Sample of a scenario pair from the stated preference experiment



Source: Authors' own elaboration

Following the pilot study, a power analysis was conducted within JMP to validate the adequacy of the design. By simulating anticipated coefficients (±1) and setting a standard significance threshold ($\alpha = 0.05$), the analysis revealed statistical power values between 0.869 and 0.907 across all attributes. These values exceed the generally accepted threshold of 0.8 (Bridges *et al.*, 2011), indicating sufficient statistical power to detect moderate effects. The final design achieved a D-efficiency score of 98.15%, indicating an optimal balance between design complexity and estimation quality.

Sample Size Determination

The ideal sample size is influenced by several factors, including the question format, the complexity of the choice tasks, the level of accuracy required for the results, the target population, the number of participants that can realistically be recruited and whether subgroup comparisons are planned (Bridges *et al.*, 2011). Traditionally, researchers have relied on heuristic guidelines to guide the sample size decisions. According to Marshall *et al.* (2010), the average sample size in conjoint

analysis studies focused on healthcare was approximately comprised between 100 and 300 respondents.

Due to the absence of a universally accepted standard for calculating minimum sample sizes in discrete choice experiments (DCEs) and to assess statistical power, we applied a commonly used rule of thumb (Orme's Rule of Thumb) suggested by de Bekker-Grob *et al*. (2015). In this study, a design comprising five attributes, a maximum of three levels ($c = 3$), 12 choice tasks per respondent ($t = 12$), and two alternatives per task ($a = 2$) was examined. Based on Orme's formula, the minimum recommended sample size to estimate main effects is:

$$n \geq \frac{1{,}000 \cdot c}{t \cdot a} = \frac{1{,}000 \cdot 3}{12 \cdot 2} = 125 \quad [1]$$

where:

n = minimum sample size

c = maximum number of levels for any attribute

t = number of tasks (choice sets) per respondent

a = number of alternatives per task (excluding "opt-out" if present)

However, since this formula ignores response variance, statistical power, and effect size, we also performed a simulation-based power analysis using the *cjpowR* package in R to increase precision, aiming for an Average Marginal Component Effect (AMCE) of 0.3, 80% statistical power, and 5% significance level ($\alpha = 0.05$). According to this analysis, a sample size of 119 respondents would be adequate.

Finally, a *simulation-based power analysis* was conducted to determine the minimum required sample size for detecting a moderate effect (AMCE = 0.3) using logistic regression, while at the same time accounting for the complexity of the experimental design, which comprised 12 choice tasks, 2 alternatives per task, and 3 levels for each attribute. According to this analysis, a sample of 80 participants would yield about 83% power, while 100 participants would provide 93% power. In DCEs, these results lend credence to the use of simulation-based techniques for more precise and design-specific sample size planning.

Participants were recruited through the CCIO Forum, expanding outreach also to CNIO and CIO networks[10], the Alan Turing Institute[11], the Futures AI Virtual Hub[12], and the Artificial Intelligence Ambassador Network[13]. To broaden engagement, the survey was also promoted via LinkedIn[14]. To

[10] Digital Health Networks. https://discourse.digitalhealth.net/
[11] The Alan Turing Institute. https://www.turing.ac.uk/
[12] AI Virtual Hub – Futures. https://future.nhs.uk/
[13] AI Ambassador Network. https://nwltraininghub.co.uk/ai-ambassador-network/
[14] Wyatt, J. C. [@jeremywyatt]. (2025, November 15). *Do we value accuracy enough when deciding which AI to purchase for the NHS?* AI can be inaccurate, biased, hard to use or fit into existing technical infrastructure… [LinkedIn

incentivise participation, respondents were given the opportunity to enter a prize draw to win one of three £50 Amazon vouchers. As of the time of writing, 42 valid responses have been collected, with a target of at least 80 participants to ensure statistical robustness.

Recruitment efforts were particularly focused on NHS professionals engaged in any phase of AI procurement, including formal decision-making, advisory, and consultative roles. Eligible participants include clinical personnel, Chief Clinical Information Officers (CCIOs), Chief Nursing Information Officers (CNIOs), Chief Information Officers (CIOs), and trainees preparing for such positions. Recruitment materials, including sample messages, infographics, and promotional content, are provided in *Appendix 3. Stages of Attribute Development (AI-CDSS Procurement Context*).

Model Specification and Analysis

The responses collected were analysed within the framework of McFadden's Random Utility Theory (RUT) (1974), which assumes that individuals make choices to maximize utility based on the attributes of the alternatives, rather than the alternatives as wholes (McFadden, 1974; Holmes & Adamowicz, 2003). RUT models are characterised as being partly observable and partly stochastic. Specifically, the observable (systematic) component reflects known attributes of the alternatives, while the unobservable component accounts for individual-specific randomness in preferences.

According to this theory, each respondent (*n*) selected the most preferred (i.e., utility-maximising) option among three alternatives across 12 choice sets. The utility (U_{nj}) that the respondent derives from alternative *j* is defined as follows:

$$U_{nj} = V_{nj} + \varepsilon_{nj} \qquad [2]$$

where V_{nj} is the observable utility and ϵ_{nj} is a random error term. With k being the number of attributes, the systematic utility is expressed as:

$$V_{nj} = \sum_{k=1}^{K} \beta_k X_{njk} \quad [3a]$$

The observable utility is modelled as a linear-in-parameters function, where β coefficients represent part-worth utilities or preference weights associated with each attribute level (Hensher *et al*., 2015):

$$V_{nj} = \beta_1 X_{nj1} + \beta_2 X_{nj2} + .. + \beta_k X_{njk} = X'_{nj}\beta \qquad [3b]$$

and the overall utility becomes:

$$U_{nj} = X'_{nj}\beta_n + \varepsilon_{nj} \qquad [4]$$

post]. LinkedIn. https://www.linkedin.com/posts/jeremy-wyatt-2ab23614_nhs-aiinhealthcare-healthcareai-activity-7392237330306220032-BPM5

where β_n captures the individual-level preference weights.

To account for preference heterogeneity among respondents (in this case, a mix of NHS staff involved in advising on AI purchasing), a Mixed Logit Model (MXL) was employed. The MXL allows for random variation in preferences across individuals by decomposing the individual-specific coefficients as:

$$\beta_n = \beta + +\eta_n \qquad [5]$$

with β as the population mean and η_n as the random deviation capturing heterogeneity. The utility function is thus expanded to:

$$U_{nj} = X'_{nj}\beta + X'_{nj}\eta_n + \varepsilon_{nj} \qquad [6]$$

The stochastic portion of utility, $X'_{nj}\eta_n + \varepsilon_{nj}$, is correlated across choice situations due to the presence of shared individual-level effects.

In this study, in which the choice model was based on five attributes of AI-enabled diagnostic systems, the utility specification for respondent *n* and alternative *j* becomes:

$$U_{nj} = \beta_0 + \beta_{1n}Usability_j + \beta_{2n}Fairness_j + \beta_{3n}Accuracy_j + \beta_{4n}Explainability + \beta_{5n}Standards_j \quad [7]$$

where β_0 represents the *Alternative Specific Constant* (ASC) associated with the opt-out option, and each $\beta_{kn} \sim N(\beta_k, \sigma_k)$, allowing the model to estimate both mean effects and standard deviations, thereby capturing variations in preferences across respondents for the given attributes.

For the purpose of statistical estimation, the conditional logit model (McFadden, 1973) was employed, a model which has wide application in the analysis of choice data. The estimation of fixed effects conditional logit models was conducted utilising R Studio with effects coding. This coding approach facilitates the interpretation of attribute-level estimates relative to the mean of other levels within the same attribute. The "no-choice" option was modelled using an additional dummy variable. The data processing began with the import of Qualtrics survey exports and the DCE design matrix. Data cleaning involved resolving formatting anomalies, converting to long format, and validating attribute-level distributions. Since alternative C (opt-out option, "Neither") introduced collinearity due to zero-coded attributes, it was explicitly modelled with an opt-out dummy and its attributes set as missing.

A conditional logit model was estimated using the *mlogit* package in R, with five attributes being used to inform the model: usability, fairness, accuracy, explainability, and adaptability to standards. The validity of the design was confirmed by diagnostics such as QR decomposition and alias matrix checks, which also minimised multicollinearity.

Findings

The survey was administered online using Qualtrics[XM15] between August and December 2025 and completed by a total of 42 professionals directly involved in the procurement of AI-based systems. Participants accessed the questionnaire through a secure, anonymized link. After initial screening, 9 responses were excluded – participants no longer holding relevant roles and departmental mismatch – resulting in a final analytic sample of 33 participants. To ensure response quality, we applied minimum completion time thresholds for the choice tasks: 15 seconds for the first scenario and at least 5 seconds for each subsequent task. All retained responses met these quality criteria. Although the minimum required sample size of 80 respondents, as calculated by the authors, has not yet been exceeded by the final sample, thereby ensuring sufficient statistical power for the empirical analyses, this preliminary analysis is based on a smaller subset (N=33) and provides an initial understanding of respondent preferences to inform ongoing data collection.

The conditional logit model revealed that clinicians place the greatest negative weight on low fairness (fairness3, $\beta = -1.67$), suggesting a strong aversion to systems perceived as biased or inequitable. In a similar vein, opaque systems (explain3, $\beta = -0.71$) and low usability (usability3, $\beta = -0.36$) were associated with lower utility; however, the effects did not attain statistical significance at the 5% level. Higher diagnostic accuracy (accuracy2, $\beta = 0.97$) had a positive influence on choice likelihood, indicating that clinicians place a value on performance, albeit within the bounds of expected clinical thresholds. While none of the coefficients attained conventional significance ($p < 0.05$), the directionality of the estimates aligns with qualitative expectations: clinicians prefer AI-CDSS tools that are accurate, explainable, fair, and usable. The preliminary results, derived from a sample of 33 NHS-affiliated professionals, provide early evidence that ethical and interpretability concerns have a significant impact on clinical procurement decisions.

[15] Qualtrics[XM]. https://www.qualtrics.com/

Table 15 – Conditional Logit Model Coefficients

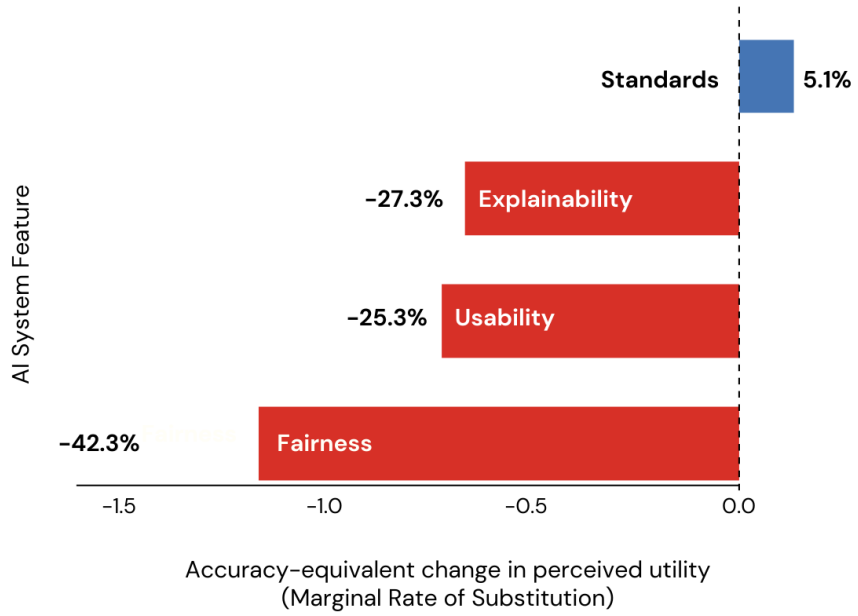
Attribute	Estimate	Std. Error	z-value	p-value	CI Lower	CI Upper
usability2	-0.14	0.40	-0.36	0.72	-0.94	0.65
usability3	-0.79	0.44	-1.8	0.07	-1.6	0.07
fairness2	-0.31	0.24	-1.3	0.19	-0.77	0.16
fairness3	-1.0	1.0	-0.99	0.32	-3.1	1.0
accuracy2	0.20	0.65	0.31	0.76	-1.1	1.5
accuracy3	-0.33	0.59	-0.56	0.58	-1.5	0.83
explain2	-0.60	0.34	-1.8	0.07	-1.3	0.06
explain3	-0.72	0.39	-1.9	0.06	-1.5	0.04
standards2	0.23	0.36	0.66	0.51	-0.46	0.93
standards3	-0.39	0.24	-1.6	0.10	-0.85	0.08

Source: our elaboration

Note: These results reflect interim analysis based on partial sample completion and will be updated in the final version of the study

Following to the methodology used by Wyatt *et al*. (2010), we calculated the Marginal Rates of Substitution (MRS) for each non-accuracy attribute using diagnostic accuracy as the baseline for comparison, to convert the utility estimates into practical trade-offs. Keeping overall perceived utility constant, the MRS values show how much more diagnostic accuracy would be needed to compensate for a one-level decline in other system features. The most important characteristic was found to be fairness: to maintain equivalent utility, a one-level drop in perceived algorithmic fairness would require a 1.3-unit (or 42%) increase in diagnostic accuracy. Similarly, a gain in accuracy of about 1 unit (≈25–27%) would be needed to compensate for losses in explainability or usability. In contrast, improved adaptability to regulatory standards yielded a minor increase in utility, equivalent to only 0.14 units (≈5%) of an accuracy improvement. Negative values displayed in the chart (Figure) clarify how much more accurate a system must be to compensate for reductions in another feature.

Figure 2 – NHS Advisors' Willingness to Trade Accuracy for Ethical and Usability Features in AI-CDSS



Source: our elaboration

Note: These results reflect interim analysis based on partial sample completion and will be updated in the final version of the study

Discussion and Conclusions

This preliminary analysis provides insightful information about how NHS professionals assess and weigh competing aspects of AI-CDSSs when advising on or making procurement decisions. Our results indicate that NHS staff preferences are more influenced by ethical and interpretability factors, especially algorithmic fairness and explainability, than by diagnostic accuracy. This trend is consistent with earlier research (Jones et al., 2023; Kovari, 2024; Giebel et al., 2025) highlighting the crucial role that trust, transparency, and usability play in the adoption of AI in healthcare. The severe utility penalty imposed on unfair systems is one of the most notable results: respondents needed a 42% increase in diagnostic accuracy to make up for a one-level drop in perceived fairness. This demonstrates that adoption cannot be guaranteed by technical performance alone as NHS procurement stakeholders appear to prioritize systems that align with ethical standards and offer clear, interpretable outputs.

To quantify these trade-offs, we computed marginal rates of substitution (MRS) following the methodology proposed by Wyatt et al. (2010). This allows us to express respondents' preferences in intuitive terms. Specifically, participants :

1. were willing to tolerate a 5% drop in accuracy if the AI system met NHS technical standards,
2. required 25% more accuracy to compensate for poor explanations and,
3. 27% more accuracy to compensate for an AI system with lower usability,
4. demanded a substantial 42% increase in accuracy to tolerate a drop in fairness (systems perceived as biased or unfair are strongly penalized).

These findings illustrate the real-world compromises or trade-offs that NHS decision-makers are willing (or unwilling) to make during AI procurement. Furthermore, even in exchange for significant performance gains, clinicians are only willing to accept modest reductions in transparency or fairness, underscoring the fact that ethical design considerations are fundamental drivers of procurement preference rather than merely incidental.

Limitations and Future Work

Notwithstanding the clarity and significance of the initial indications, it must be acknowledged that these findings remain preliminary. This is due to the fact that the analysis is grounded in a partial dataset comprising 33 participants. Larger and more varied samples will be needed to ensure statistical robustness and generalizability, even though the initial trends are promising. At least 80 NHS professionals, including those in formal and advisory procurement roles (CCIOs, CNIOs, CIOs, clinical staff, and trainees), will be included in the sample through ongoing recruitment. Heterogeneity in preferences across professional groups and clinical domains will also be investigated in future research, potentially through latent class analysis or mixed logit models.

Strengths of the Research

A key strength of this study lies in its innovative application of McFadden's Random Utility Theory (RUT) to model NHS decision-making in AI procurement, a framework that captures both systematic and individual-specific components of utility. The experimental design also included data quality controls: a minimum completion time threshold of 15 seconds for the first task and 5 seconds for each subsequent task ensured that respondents thoughtfully engaged with each choice scenario. The use of marginal rate of substitution (MRS) analysis further enhances the interpretability of results by expressing trade-offs in policy-relevant, intuitive terms.

Research Implications

From a methodological standpoint, this study shows that NHS advisors and decision-makers' operational and ethical priorities in AI procurement can be successfully captured by discrete choice experiments (DCEs). The method offers a transparent and quantitative way to incorporate stakeholder preferences into decision frameworks, and it can be repeated to investigate other technology evaluation settings. To further understand the factors influencing AI acceptance, researchers can expand on this work by adding behavioural and contextual factors (such as prior AI experience or organizational culture).

Practical and Policy Implications

The real-world significance of this research is highlighted by its alignment with current NHS policy initiatives, such as the upcoming update of the NHS AI Buyers Guide (originally published in 2020).

NHS England policy consultants have confirmed that the research team's findings will influence the revised procurement guidelines. The open consultation (closing 30 November 2025) and subsequent policy submission (19 December 2025) create a timely opportunity for direct policy integration. By ensuring that fairness, explainability, and usability are given equal weight with performance metrics in AI purchasing decisions, these insights can aid in the development of ethical procurement standards. These findings will be translated into practical recommendations for the responsible, open adoption of AI throughout the healthcare system with continued engagement with NHS stakeholders, especially through the upcoming CCIO Conference. These next steps will help bridge the gap between empirical findings and actionable recommendations for ethical, effective AI adoption in the NHS.

In summary, this study reveals a decisive trend: when evaluating an AI system, NHS professionals value ethical alignment and interpretability as crucial as the system's accuracy. Despite the preliminary nature of these results, they offer early indications that the adoption of AI in clinical settings may be contingent on effective design. As data collection expands and policy collaboration deepens, this research is well-positioned to contribute directly to the next generation of NHS AI procurement, policy, and practice.

References

Adler-Milstein, J., Aggarwal, N., Ahmed, M., Castner, J., Evans, B. J., Gonzalez, A. A., James, C. A., Lin, S., Mandl, K. D., Matheny, M. E., Sendak, M. P., Shachar, C., & Williams, A. (2022). Meeting the Moment: Addressing Barriers and Facilitating Clinical Adoption of Artificial Intelligence in Medical Diagnosis. *NAM perspectives*, *2022*, 10.31478/202209c. https://doi.org/10.31478/202209c

Afful-Dadzie, E., & Afful-Dadzie, A. (2021). Online health consumer behaviour: What informs user decisions on information quality? Computers in Human Behavior Reports, 3, 100064. https://doi.org/10.1016/j.chbr.2021.100064

Afful-Dadzie, E., & Egala, S. B. (2022). Medical practitioners' decision making on quality of online medical information: A consumption values theory analysis. *Health Policy and Technology*, 11(4), 100685. https://doi.org/10.1016/j.hlpt.2022.100685

Akinc, D., & Vandebroek, M. (2017). Comparing the performances of maximum simulated likelihood and hierarchical Bayesian estimation for mixed logit models. *International Choice Modelling Conference* 2017. https://doi.org/10.2139/ssrn.3052293

Alshehhi, K., Cheaitou, A., Rashid, H. (2023). Fuzzy Failure Modes Effect and Criticality Analysis of the Procurement Process of Artificial Intelligent Systems/Services. *International Journal of Advanced Computer Science and Applications*, *14*(10). http://dx.doi.org/10.14569/IJACSA.2023.0141060

Badnjevic, A., Avdihodžić, H., & Gurbeta Pokvic, L. (2021). Artificial Intelligence in Medical Devices: Past, Present and Future. *Psychiatria Danubina*, *33*, 101–106. https://doi.org/10.5005/sar-1-1-2-101

Bai, E., Zhang, Z., Xu, Y., Luo, X., & Adelgais, K. (2024). Enhancing Prehospital Decision-Making: Exploring User Needs and Design Considerations for Clinical Decision Support Systems. *Research square*, rs.3.rs-5206138. https://doi.org/10.21203/rs.3.rs-5206138/v1

Bankuoru Egala, S., & Liang, D. (2024). Algorithm aversion to mobile clinical decision support among clinicians: a choice-based conjoint analysis. *European Journal of Information Systems*, *33*(6), 1016-1032. https://doi.org/10.1080/0960085X.2023.2251927

Ben-Akiva, M., McFadden, D., & Train, K. (2019). Foundations of Stated Preference Elicitation: Consumer Behavior and Choice-based Conjoint Analysis. *Foundations and Trends® in Econometrics*, *10*, 1–144. https://doi.org/10.1561/0800000036

Benlian, A. (2011). Is traditional, open-source, or on-demand first choice? Developing an AHP-based framework for the comparison of different software models in office suites selection. European Journal of Information Systems, 20(5), 542–559. https://doi.org/10.1057/ejis.2011.14

Bertl, M., Ross, P., & Draheim, D. (2023). Systematic AI Support for Decision-Making in the Healthcare Sector: Obstacles and Success Factors. *Health Policy and Technology*, *12*(3), 100748. https://doi.org/10.1016/j.hlpt.2023.100748

Bonakdeh, E. S., Sohal, A., Moghadam, V. K., Rajabkhah, K., Prajogo, D., Melder, A., Nguyen, Q., Bingham, G., & Tong, E. (2025). The facilitating and hindering factors of pre-implementation phase in establishment of clinical decision support systems: A systematic review and meta synthesis. *Health Policy and Technology*, *14*(3), 101014. https://doi.org/10.1016/j.hlpt.2025.101014

Bridges, J. F., Hauber, A. B., Marshall, D., Lloyd, A., Prosser, L. A., Regier, D. A., Johnson, F. R., & Mauskopf, J. (2011). Conjoint analysis applications in health-a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*, *14*(4), 403–413. https://doi.org/10.1016/j.jval.2010.11.013

Broekhuizen, T., Dekker, H., de Faria, P., Firk, S., Nguyen, D. K., & Sofka, W. (2023). AI for managing open innovation: Opportunities, challenges, and a research agenda. *Journal of Business Research*, *167*, 114196. https://doi.org/10.1016/j.jbusres.2023.114196

Chada, B. V., & Summers, L. (2022). AI in the NHS: a framework for adoption. *Future Healthcare Journal*, *9*(3), 313–316. https://doi.org/10.7861/fhj.2022-0068

Coast, J., Al-Janabi, H., Sutton, E. J., Horrocks, S. A., Vosper, A. J., Swancutt, D. R., & Flynn, T. N. (2012). Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health economics*, 21(6), 730–741. https://doi.org/10.1002/hec.1739

Cunningham, C. E., Deal, K., & Chen, Y. (2010). Adaptive Choice-Based Conjoint Analysis. *The Patient: Patient-Centered Outcomes Research*, *3*(4), 257–273. https://doi.org/10.2165/11537870-000000000-00000

Danner, M., Vennedey, V., Hiligsmann, M., Fauser, S., Gross, C., & Stock, S. (2017). Comparing analytic hierarchy process and discrete-choice experiment to elicit patient preferences for treatment characteristics in age-related macular degeneration. *Value in Health*, 20(8), 1166–1173. https://doi.org/10.1016/j.jval.2017.04.022

Daye, D., Wiggins, W. F., Lungren, M. P., Alkasab, T., Kottler, N., Allen, B., Roth, C. J., Bizzo, B. C., Durniak, K., Brink, J. A., Larson, D. B., Dreyer, K. J., & Langlotz, C. P. (2022). Implementation of Clinical Artificial Intelligence in Radiology: Who Decides and How?. *Radiology*, *305*(3), 555–563. https://doi.org/10.1148/radiol.212151

de Bekker-Grob, E. W., Ryan, M., & Gerard, K. (2012). Discrete choice experiments in health economics: A review of the literature. *Health Economics, 21*(2), 145–172.

de Hond, A. A. H., Leeuwenberg, A. M., Hooft, L., Kant, I. M. J., Nijman, S. W. J., van Os, H. J. A., Aardoom, J. J., Debray, T. P. A., Schuit, E., van Smeden, M., Reitsma, J. B., Steyerberg, E. W., Chavannes, N. H., & Moons, K. G. M. (2022). Guidelines and quality criteria for artificial intelligence-based prediction models in healthcare: a scoping review. *NPJ digital medicine*, *5*(1), 2. https://doi.org/10.1038/s41746-021-00549-7

Dellink, R. (2023). *Barriers to the adoption of artificial intelligence in medical diagnosis: Procurement perspective* (public). http://essay.utwente.nl/96610/

Dingel, J., Kleine, A. K., Cecil, J., Sigl, A. L., Lermer, E., & Gaube, S. (2024). Predictors of Health Care Practitioners' Intention to Use AI-Enabled Clinical Decision Support Systems: Meta-Analysis Based on the Unified Theory of Acceptance and Use of Technology. Journal of medical Internet research, 26, e57224. https://doi.org/10.2196/57224

Dong, D., Ozdemir, S., Bee, Y. M., Toh, S.-A., Bilger, M., & Finkelstein, E. (2016). Measuring high-risk patients' preferences for pharmacogenetic testing to reduce severe adverse drug reactions: A discrete choice experiment. *Value in Health, 19*(6), 767–775.

Esan, O., Uzozie, O., Onaghinor, O., Etukudoh, E., Omisola, J., & Pub, A. (2023). Agile Procurement Management in the Digital Age: A Framework for Data-Driven Vendor Risk and Compliance Assessment. *International Journal of Management and Organizational Research*, 02, 320–327.

Evans, T., Ahmad, O., Alderman, J., Bailey, G., Bannister, P., Barlow, N., Davison, N., Isaac, A., Kale, A., Macdonald, T., Malik, Q., Shelmerdine, S., Hogg, H., & Denniston, A. (2025). The role of procurement frameworks in responsible AI innovation in the National Health Service: A multi-stakeholder perspective. *Frontiers in health services*, *5*, 1608087. https://doi.org/10.3389/frhs.2025.1608087

Finkelstein, J., Gabriel, A., Schmer, S., Truong, T. T., & Dunn, A. (2024). Identifying Facilitators and Barriers to Implementation of AI-Assisted Clinical Decision Support in an Electronic Health Record System. *Journal of medical systems*, *48*(1), 89. https://doi.org/10.1007/s10916-024-02104-9

Fujimori, R., Liu, K., Soeno, S., Naraba, H., Ogura, K., Hara, K., Sonoo, T., Ogura, T., Nakamura, K., & Goto, T. (2022). Acceptance, Barriers, and Facilitators to Implementing Artificial Intelligence-Based Decision Support Systems in Emergency Departments: Quantitative and Qualitative Evaluation. *JMIR formative research*, *6*(6), e36501. https://doi.org/10.2196/36501

Gama, F., Tyskbo, D., Nygren, J., Barlow, J., Reed, J., & Svedberg, P. (2022). *Implementation frameworks for artificial intelligence translation into health care practice: scoping review*. Journal of Medical Internet Research, 24(1): e32215. https://doi.org/10.2196/32215

Giebel, G. D., Raszke, P., Nowak, H., Palmowski, L., Adamzik, M., Heinz, P., Tokic, M., Timmesfeld, N., Brunkhorst, F., Wasem, J., & Blase, N. (2025). Problems and Barriers Related to the Use of AI-Based Clinical Decision Support Systems: Interview Study. *J Med Internet Res*, *27*, e63377. https://doi.org/10.2196/63377

Goldfarb A, Teodoridis F. Why is AI adoption in health care lagging? Washington DC: Brookings Institution; 2022.

Grammatopoulos, D., Braybrook, E., & Anderson, N. (2024). *THE UK LABORATORY DIAGNOSTICS LANDSCAPE REPORT University of Warwick and UHCW NHS Trust*. https://doi.org/10.5281/zenodo.13970649

Haaijer, R., Kamakura, W. A., & Wedel, M. (2001). The 'no-choice' alternative in conjoint choice experiments. *International Journal of Market Research*, 43.

Hahn, G. J., Shapiro, S. S., & General Electric Company Research and Development Center. (1966). *A catalog and computer program for the design and analysis of orthogonal symmetric and asymmetric fractional factorial experiments*. General Electric Research and Development Center; WorldCat.

Hauser, J. R., & Toubia, O. (2005). The impact of utility balance and endogeneity in conjoint analysis. *Marketing Science, 24*(3), 498–507.

Helm, R., Steiner, M., Scholl, A., & Manthey, L. (2008). A comparative empirical study on common methods for measuring preferences. *International Journal of Management & Decision Making*, 9(3), 242–265. https://doi.org/10.1504/IJMDM.2008.017408

Helter, T. M., & Boehler, C. E. H. (2016). Developing attributes for discrete choice experiments in health: A systematic literature review and case study of alcohol misuse interventions. *Journal of Substance Use*, 21(6), 662–670.

Hensher, D. A., Rose, J. M., & Greene, W. H. (2015). *Applied choice analysis*. Cambridge university press.

HSSIB. *Electronic patient record (EPR) systems – thematic review.* Published 5 March 2025. https://www.hssib.org.uk/patient-safety-investigations/electronic-patient-record-epr-systems-thematic-review/

Huls, S. P. I., van Exel, J., & de Bekker-Grob, E. W. (2023). An attempt to decrease social desirability bias: The effect of cheap talk mitigation on internal and external validity of discrete choice experiments. *Food Quality and Preference*, *111*, 104986. https://doi.org/10.1016/j.foodqual.2023.104986

Jones, C., Thornton, J., & Wyatt, J. C. (2023). Artificial intelligence and clinical decision support: clinicians' perspectives on trust, trustworthiness, and liability. *Medical law review*, *31*(4), 501–520. https://doi.org/10.1093/medlaw/fwad013

Karakaya, F., & Awasthi, A. (2015). Robustness and sensitivity of conjoint analysis versus multiple linear regression analysis. International *Journal of Data Analysis Techniques and Strategies*, 7(2), 121–136. https://doi.org/10.1504/IJDATS.2014.062461

Kawamoto, K., Houlihan, C. A., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ (Clinical research ed.)*, *330*(7494), 765. https://doi.org/10.1136/bmj.38398.500764.8F

Kessels, R., Jones, B., & Goos, P. (2011). Bayesian optimal designs for discrete choice experiments with partial profiles. *Journal of Choice Modelling, 4*(3), 52–74.

Kim, J. Y., et al. (2023). *Organizational Governance of Emerging Technologies: AI Adoption in Healthcare*. FAccT. https://doi.org/10.1145/3593013.3594089

Kočo, L., Siebers, C. C. N., Schlooz, M., Meeuwis, C., Oldenburg, H. S. A., Prokop, M., & Mann, R. M. (2024). The Facilitators and Barriers of the Implementation of a Clinical Decision Support System for Breast Cancer Multidisciplinary Team Meetings-An Interview Study. *Cancers*, *16*(2), 401. https://doi.org/10.3390/cancers16020401

Kovari, A. (2024). AI for Decision Support: Balancing Accuracy, Transparency, and Trust Across Sectors. *Information*, *15*(11), 725. https://doi.org/10.3390/info15110725

Lancsar, E., & Louviere, J. (2008). Conducting discrete choice experiments to inform healthcare decision making: A user's guide. *Pharmacoeconomics, 26*(8), 661–677.

Latt, P. M., Soe, N. N., King, A. J., Lee, D., Phillips, T. R., Xu, X., Chow, E. P. F., Fairley, C. K., Zhang, L., & Ong, J. J. (2024). Preferences for attributes of an artificial intelligence-based risk assessment tool for HIV and sexually transmitted infections: A discrete choice experiment. *BMC Public Health*, *24*(1), 3236. https://doi.org/10.1186/s12889-024-20688-2

Lenk, P. J., DeSarbo, W. S., Green, P. E., & Young, M. R. (1996). Hierarchical Bayes conjoint analysis: Recovery of partworth heterogeneity from reduced experimental designs. *Marketing Science*, 15(2), 173–191. https://doi.org/10.1287/mksc.15.2.173

Louder, A. M., Singh, A., Saverno, K., Cappelleri, J. C., Aten, A. J., Koenig, A. S., & Pasquale, M. K. (2016). Patient Preferences Regarding Rheumatoid Arthritis Therapies: A Conjoint Analysis. *American health & drug benefits*, 9(2), 84–93.

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: Analysis and application*. Cambridge University Press.

Louviere, J., Hensher, D., & Swait, J. (2007). Conjoint preference elicitation methods in the broader context of random utility theory preference elicitation methods. In A. Gustafsson, A. Herrmann, & F. Huber (Eds.), *Conjoint measurement: Methods and applications* (pp. 167–197). Springer. https://doi.org/10.1007/978-3-540-71404-0_10

Mangham, L. J., Hanson, K., & McPake, B. (2009). How to do (or not to do)… Designing a discrete choice experiment for application in a low-income country. *Health Policy and Planning, 24*(2), 151–158.

Marshall, D., Bridges, J. F., Hauber, B., Cameron, R., Donnalley, L., Fyie, K., & Johnson, F. R. (2010). Conjoint Analysis Applications in Health - How are Studies being Designed and Reported?: An Update on Current Practice in the Published Literature between 2005 and 2008. *The patient*, *3*(4), 249–256. https://doi.org/10.2165/11539650-000000000-00000

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105–142). Academic Press.

Medicines and Healthcare products Regulatory Agency (MHRA). (2023). *Software and artificial intelligence (AI) as a medical device: Change programme roadmap.* UK Government. Retrieved from https://www.gov.uk/guidance/software-and-ai-as-a-medical-device-change-programme

Meehan, J., Menzies, L., & Michaelides, R. (2017). The long shadow of public policy; Barriers to a value-based approach in healthcare procurement. *Special issue of the 25th annual IPSERA conference, Dortmund, Germany, 2016Purchasing & Supply Management: From efficiency to effectiveness in an integrated supply chain*, *23*(4), 229–241. https://doi.org/10.1016/j.pursup.2017.05.003

Mele N. L. (2008). Conjoint analysis: using a market-based research model for healthcare decision making. *Nursing research*, *57*(3), 220–224. https://doi.org/10.1097/01.NNR.0000319499.52122.d2

Mele, N. L. (2008). Conjoint Analysis: Using a Market-Based Research Model for Healthcare Decision Making. *Nursing Research*, *57*(3). https://journals.lww.com/nursingresearchonline/fulltext/2008/05000/conjoint_analysis__using_a_market_based_research.13.aspx

Miguel, F. S., Ryan, M., & Amaya-Amaya, M. (2005). 'Irrational' stated preferences: A quantitative and qualitative investigation. *Health Economics*, 14(3), 307–322.

Morandini, S., Fraboni, F., Balatti, E., Hackmann, A., Brendel, H., Puzzo, G., Volpi, L., Giusino, D., De Angelis, M., & Pietrantoni, L. (2023). *Assessing the Transparency and Explainability of AI Algorithms in Planning and Scheduling tools: A Review of the Literature*. https://doi.org/10.54941/ahfe1004068

Muehlematter, U. J., Daniore, P., & Vokinger, K. N. (2021). *Approval of AI/ML-based medical devices in the USA and Europe (2015–20): A comparative analysis*. The Lancet Digital Health, 3(3), e195–e203. https://doi.org/10.1016/S2589-7500(20)30292-2

Mühlbacher, A., & Johnson, F. R. (2016). Choice Experiments to Quantify Preferences for Health and Healthcare: State of the Practice. *Applied Health Economics and Health Policy*, *14*(3), 253–266. https://doi.org/10.1007/s40258-016-0232-7

National Institute for Health and Care Excellence (NICE). (2022). *Evidence standards framework for digital health technologies.* NICE. Retrieved from https://www.nice.org.uk/about/what-we-do/our-programmes/evidence-standards-framework-for-digital-health-technologies

NHS AI Lab. (2023). *AI assurance toolkit: Guidelines for safe, effective, and ethical AI deployment in health and care.* NHS England. Retrieved from https://transform.england.nhs.uk/ai-lab/ai-assurance/

NHSX. (2021). *A buyer's guide to AI in health and care.* NHS England. Retrieved November 13, 2025, from https://webarchive.nationalarchives.gov.uk/ukgwa/20230504130145mp_/https://transform.england.nhs.uk/media/documents/NHSX_A_Buyers_Guide_to_AI_in_Health_and_Care.pdf

Office for Artificial Intelligence, Cabinet Office, & World Economic Forum Centre for the Fourth Industrial Revolution. (2020). *Guidelines for AI procurement: A summary of best practices addressing specific challenges of acquiring Artificial Intelligence in the public sector*. Crown Copyright. https://www.gov.uk/government/publications/guidelines-for-ai-procurement

Peek, N., Capurro, D., Rozova, V., & van der Veer, S. N. (2024). Bridging the Gap: Challenges and Strategies for the Implementation of Artificial Intelligence-based Clinical Decision Support Systems in Clinical Practice. *Yearbook of medical informatics*, *33*(1), 103–114. https://doi.org/10.1055/s-0044-1800729

Petersson, L., Larsson, I., Nygren, J. M., Nilsen, P., Neher, M., Reed, J. E., Tyskbo, D., & Svedberg, P. (2022). Challenges to implementing artificial intelligence in healthcare: a qualitative interview study with healthcare leaders in Sweden. *BMC health services research*, *22*(1), 850. https://doi.org/10.1186/s12913-022-08215-8

Ploug, T., Sundby, A., Moeslund, T. B., & Holm, S. (2021). Population Preferences for Performance and Explainability of Artificial Intelligence in Health Care: Choice-Based Conjoint Survey. *Journal of medical Internet research*, *23*(12), e26611. https://doi.org/10.2196/26611

Raghavarao, D., Wiley, J. B., & Chitturi, P. (2010). *Choice-based conjoint analysis: Models and designs*. CRC Press. https://doi.org/10.1201/9781420099973

Rao, V. R. (2014). Applied conjoint analysis. Springer. https://doi.org/10.1007/978-3-540-87753-0

Richard, P. J., Coltman, T. R., & Keating, B. W. (2012). Designing IS service strategy: An information acceleration approach. *European Journal of Information Systems*, 21(1), 87–98. https://doi.org/10.1057/ejis.2010.62

Roosen, J., & Marette, S. (2011). Making the 'right' choice based on experiments: Regulatory decisions for food and health. European Review of Agricultural Economics, 38(3), 361–381. https://doi.org/10.1093/erae/jbr026

Roßnagel, H., Zibuschka, J., Hinz, O., & Muntermann, J. (2014). Users' willingness to pay for web identity management systems. European Journal of Information Systems, 23(1), 36–50. https://doi.org/10.1057/ejis.2013.33

Ryan, M., Gerard, K., & Amaya-Amaya, M. (2007). *Using discrete choice experiments to value health and health care*. Springer Science & Business Media.

Schmidt, K., Babac, A., Pauer, F., Damm, K., & von der Schulenburg, J. (2016). Measuring patients' priorities using the analytic hierarchy process in comparison with best-worst scaling and rating cards: Methodological aspects and ranking tasks. *Health Economics Review*, 6(1), 1–11. https://doi.org/10.1186/s13561-016-0130-6

Schouten, B. C., Cox, A., Duran, G., Kerremans, K., Banning, L. K., Lahdidioui, A., van den Muijsenbergh, M., Schinkel, S., Sungur, H., Suurmond, J., Zendedel, R., & Krystallidou, D. (2020). Mitigating language and cultural barriers in healthcare communication: Toward a holistic approach. Patient education and counseling, S0738-3991(20)30242-1. Advance online publication. https://doi.org/10.1016/j.pec.2020.05.001

Sharma, S. K., Dwivedi, Y. K., Misra, S. K., & Rana, N. P. (2023). Conjoint analysis of blockchain adoption challenges in government. *Journal of Computer Information Systems*, 00(0), 1–14. https://doi.org/10.1080/08874417.2023.2185552

Stogiannos, N., Malik, R., Kumar, A., Barnes, A., Pogose, M., Harvey, H., McEntee, M. F., & Malamateniou, C. (2023). Black box no more: a scoping review of AI governance frameworks to guide procurement and adoption of AI in medical imaging and radiotherapy in the UK. *The British journal of radiology*, *96*(1152), 20221157. https://doi.org/10.1259/bjr.20221157

Tricco, A.C., Lillie, E., Zarin, W., O'Brien, K.K., Colquhoun, H., Levac, D., Moher, D., Peters, M.D., Horsley, T., Weeks, L., Hempel, S., Akl, E.A., Chang, C., McGowan, J., Stewart, L., Hartling, L., Aldcroft, A., Wilson, M.G., Garritty, C., Lewin, S., Godfrey, C.M., Macdonald, M.T., Langlois, E.V., Soares-Weiser, K., Moriarty, J., Clifford, T., Tunçalp, O. and Straus, S.E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. *Annals of Internal Medicine*, *169*(7), 467–473. https://doi.org/10.7326/M18-0850

Uravirta, J. (2023). Factors influencing machine learning application procurement in healthcare organisations. Master's Thesis. Metropolia University of Applied Sciences. https://www.theseus.fi/handle/10024/808835

Van Houtven, G., Johnson, F. R., Kilambi, V., & Hauber, A. B. (2011). Eliciting benefit-risk preferences and probability-weighted utility using choice-format conjoint analysis. *Medical Decision Making*, 31(3), 469–480. https://doi.org/10.1177/0272989X10386116

Viney, R., Lancsar, E., & Louviere, J. (2002). Discrete choice experiments to measure consumer preferences for health and healthcare. *Expert Review of Pharmacoeconomics & Outcomes Research, 2*(4), 319–326.

von Wedel, P., & Hagist, C. (2018). Die Digitalisierung der Arzt-Patienten-Beziehung in Deutschland: Ein Discrete Choice Experiment zur Analyse der Patientenpräferenzen bezüglich digitaler Gesundheitsleistungen. *Gesundheitsökonomie & Qualitätsmanagement, 23*(3), 142–149.

Wang, D., Wang, L., Zhang, Z., Wang, D., Zhu, H., Gao, Y., Fan, X., & Tian, F. (2021). "Brilliant AI Doctor" in Rural Clinics: Challenges in AI-Powered Clinical Decision Support System Deployment. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. https://doi.org/10.1145/3411764.3445432

Woode, M. E., De Silva Perera, U., Degeling, C., Aquino, Y. S. J., Houssami, N., Carter, S. M., & Chen, G. (2025). Preferences for the Use of Artificial Intelligence for Breast Cancer Screening in Australia: A Discrete Choice Experiment. *The Patient - Patient-Centered Outcomes Research*. https://doi.org/10.1007/s40271-025-00742-w

Wu, E., Wu, K., Daneshjou, R., Ouyang, D., Ho, D. E., & Zou, J. (2021). How medical AI devices are evaluated: limitations and recommendations from an analysis of FDA approvals. *Nature medicine*, *27*(4), 582–584. https://doi.org/10.1038/s41591-021-01312-x

Wyatt, J. C., Batley, R. P., & Keen, J. (2010). GP preferences for information systems: conjoint analysis of speed, reliability, access and users. *Journal of evaluation in clinical practice*, *16*(5), 911–915. https://doi.org/10.1111/j.1365-2753.2009.01217.x

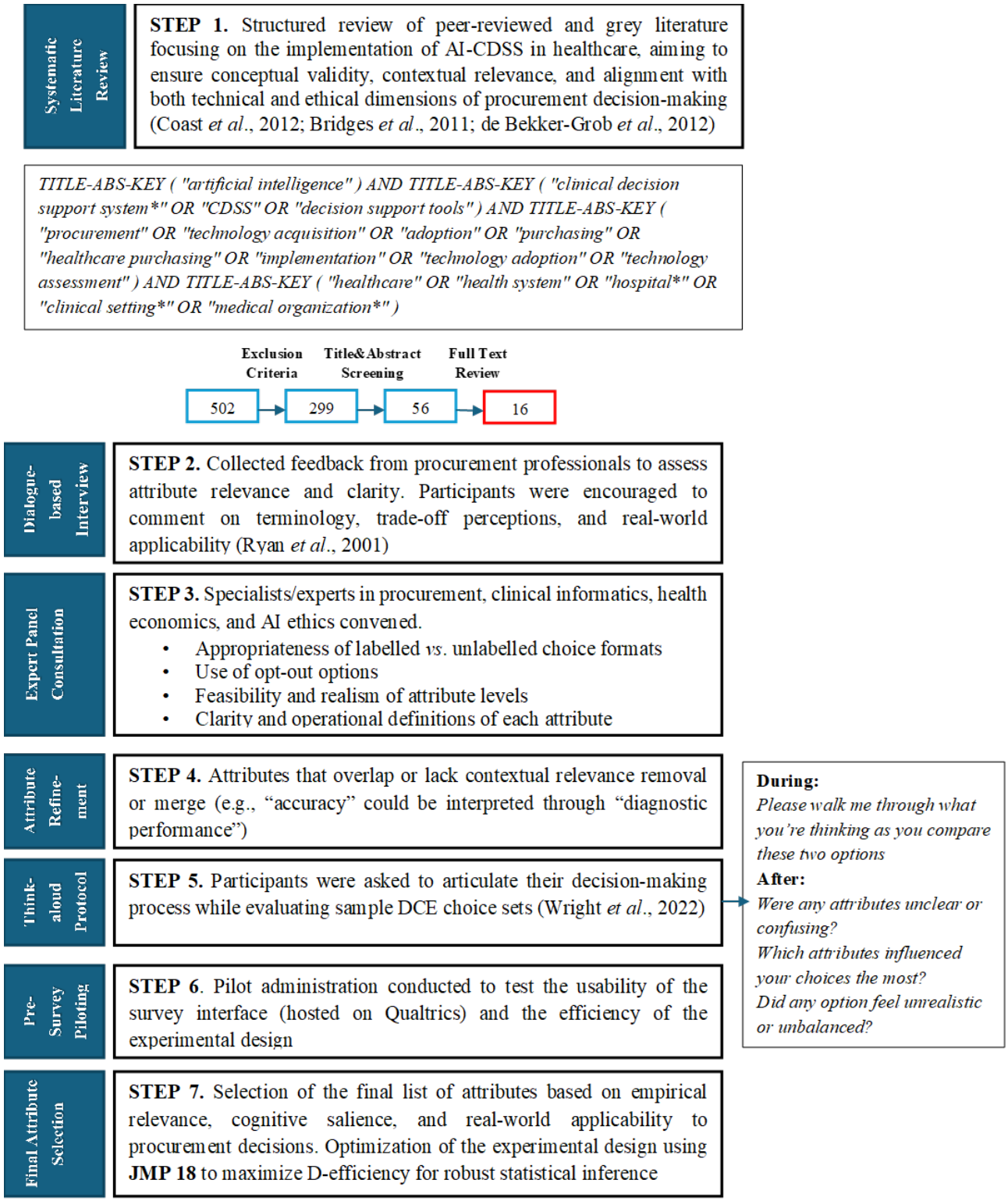
Yang, J., Luo, B., Zhao, C., & Zhang, H. (2022). Artificial intelligence healthcare service resources adoption by medical institutions based on TOE framework. *Digital health*, *8*, 20552076221126034. https://doi.org/10.1177/20552076221126034

Yeh, C., Hartmann, M., & Langen, N. (2020). The role of trust in explaining food choice: Combining choice experiment and attribute best–worst scaling. *Foods*, 9(1), 45. https://doi.org/10.3390/foods9010045

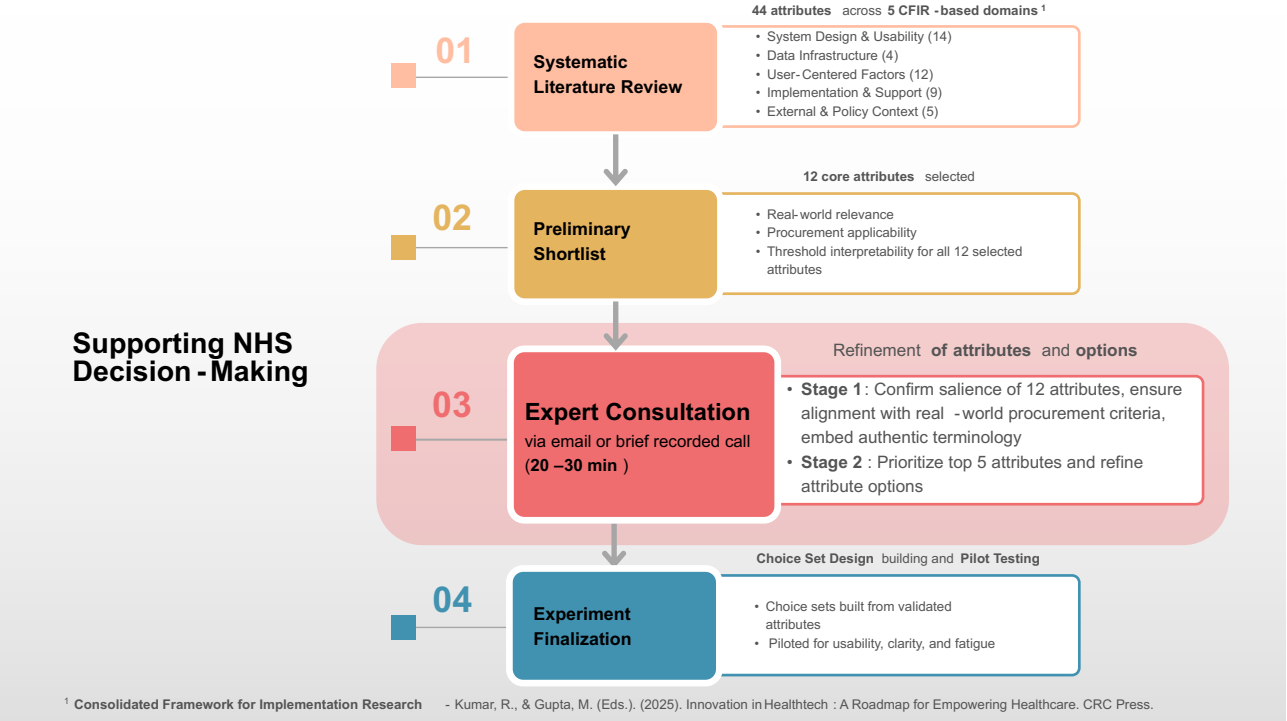
Zhang, J., Wang, J., Zhang, J., Xia, X., Zhou, Z., Zhou, X., & Wu, Y. (2025). Young Adult Perspectives on Artificial Intelligence-Based Medication Counseling in China: Discrete Choice Experiment. *Journal of medical Internet research*, *27*, e67744. https://doi.org/10.2196/67744

Zhang, J., Whebell, S., Gallifant, J., Budhdeo, S., Mattie, H., Lertvittayakumjorn, P., et al. (2022). *An interactive dashboard to track themes, development maturity, and global equity in clinical artificial intelligence research*. Lancet Digital Health, 4(4), e212–e213. https://doi.org/10.1016/S2589-7500(22)00032-2

Appendix 3. Stages of Attribute Development (AI-CDSS Procurement Context)



Attributes Refinement Process



Selected Attributes & Levels

Attr. N.	Attribute Name	Description	1st Level	2nd Level	3rd Level
1	**System Usability/ Workflow Integration**	Degree to which the AI-CDSS is intuitive and supports clinical tasks with minimal disruption	≤2 clicks, integrates with clinical pathways	3–5 clicks, occasional override	>5 clicks, frequent overrides/disruption to tasks
2	**Algorithm Update Frequency**	Frequency with which the algorithm is updated to reflect new data, clinical guidelines, or system learning	No updates	Annual Updates	Real-time retraining
3	**Algorithmic Fairness/ Bias Mitigation**	Extent to which the AI-CDSS performs equitably across diverse patient groups	Validated across multiple demographic groups	Tested in some populations, but limited generalizability	Unknown or unreported bias levels
4	**Data Processing & Privacy Architecture**	Where and how patient data is processed (local vs. cloud)	Local-only processing	Hybrid (Local+Cloud)	Full Cloud Processing
5	**Diagnostic Accuracy**	Credibility and performance of the system's predictions, based on source and quality of validation	<90% accuracy or based only on internal developer validation	90–94% accuracy validated in routine clinical settings, without peer-reviewed evidence	≥95% accuracy, validated in peer-reviewed studies

6	**Clinical Impact Scope**	Effect to which AI-CDSS affects patient outcomes at scale, combining both the proportion of patients impacted and the severity of the addressed clinical problems	Impacts <5% of patients with low clinical relevance (e.g., minor workflow gains)	Impacts 5–20% of patients or reduces diagnostic errors/complications	Impacts >20% of patients or prevents severe outcomes (e.g., mortality)
7	**Explainability**	Clarity and transparency of the AI-CDSS decision-making process in ways clinicians can interpret, trust, and defend professionally	Highly explainable with interpretable models and transparent logic	Moderate explainability; partial transparency or simplified summaries	Opaque system with no understandable rationale for outputs
8	**Cost (per patient use)**	Cost of using the AI-CDSS on a per-patient basis	£5-10 per use	£11-20 per use	£21-40 per use
9	**Technical Support**	Availability and responsiveness of technical support during implementation	No dedicated support; delayed response to issues	Support available during business hours; standard resolution time	24/7 expert support with rapid response times and dedicated contact
10	**Vendor Reliability**	Level of vendor commitment to maintenance, training, and system updates	Limited post-sale engagement or responsiveness	Moderate engagement; periodic updates and guidance	Proactive vendor with full lifecycle support and user training
11	**Ease of Adaptability to Standards**	Ability of the AI-CDSS to comply with and adapt to evolving NHS	Fully certified + adaptable to NHS/AI Act evolution (e.g., can update under regulation change)	Meets current standards (UKCA marked; MHRA standard; NICE ESF) but inflexible	Legacy certifications or unclear liability in case of error
12	**Implementation Complexity**	Extent of technical efforts, resources and configurations to deploy the AI-CDSS within existing NHS digital systems (EHR integration, API standardization etc.)	Plug-and-play integration using standard APIs	Moderate configuration and alignment with existing data structures	Major development needed: custom APIs, data translators, manual data entry

Concluding Remarks

This doctoral study combined decision sciences, behavioural economics, innovation management, and digital health policy to examine the complex issue of AI adoption and governance in healthcare from an integrated perspective. The work approached AI not merely as a technological innovation, but as a complex socio-technical and organizational process shaped by institutional frameworks, behavioural dynamics, and managerial decisions.

From an empirical standpoint, the five studies included in this thesis work together to improve our understanding of how economic actors – clinicians, organizations, and policymakers – evaluate, trust, and procure AI-based clinical decision support systems (AI-CDSSs). By applying methods ranging from systematic literature analysis, thematic content analysis, patent network mapping, behavioural modelling, and discrete choice experiments, the research has provided evidence across multiple levels of decision-making.

The first study conducted a comprehensive systematic review of the literature aimed at mapping the intellectual landscape of AI adoption in healthcare. The study's application of Prospect Theory to the literature corpus demonstrated that cognitive biases, clinicians' aversion to risk, and uncertainty continue to be significant obstacles to the integration of AI in clinical decision-making. Although AI's potential to improve diagnostic efficiency and accuracy has been amply demonstrated, the review highlighted a widespread disregard for organizational, ethical, and legal considerations, which frames the ensuing empirical studies.

Building on these insights, the second study investigated the institutional and regulatory discourse related to AI within the European context. Through a qualitative content analysis of official EU communications and policy documents (2018-2025), this study emphasised the changing definition of "trustworthy AI" within the European governance framework. The study outlined the three main policy narratives that support the proposed AI Act: risk classification, human oversight and conformity assessment. It emphasised the need for standardised governance models that strike a balance between accountability and innovation.

Through a network-based and patentometric analysis of over 8,000 AI-healthcare patents submitted globally between 2014 and 2024, the third study offered a macro-level analysis of innovation dynamics. The results showed that the industry's innovation is concentrated in a small number of highly interconnected technological ecosystems dominated by *medical imaging*, *diagnostics*, and *predictive analytics*. Network centrality measures were found to reveal the presence of high-value patents within collaborative settings, thereby confirming that inter-organizational knowledge flows and technological complementarities – predominant themes in the larger literature on innovation management – are essential to digital health innovation capacity.

The fourth study examined the contextual and psychological determinants of AI adoption by conducting a cross-sectional survey among medical professionals in Italy and the UK. The analysis revealed that clinicians' behavioural intentions toward AI-based Clinical Decision Support Systems are significantly shaped by their perceptions of the systems' usability, risk, and usefulness. Furthermore, institutional factors including perceived liability protection, regulatory clarity, and digital literacy were shown to moderate these associations. These results imply that managerial and organizational capacities to promote trust, offer sufficient training, and coordinate technology integration with institutional safeguards are essential for the successful adoption of AI in healthcare.

The final study (*ongoing*) provided important information about how NHS staff evaluate conflicting features of AI-CDSSs during procurement decisions. In particular, the preliminary results show that clinicians' preferences are dominated by fairness and explainability, surpassing diagnostic accuracy, consistently with previous research highlighting ethical and usability issues in healthcare AI adoption. By computing marginal rates of substitution (MRS) to quantify the trade-offs between different systems' features, NHS advisors demonstrated that they apply a severe utility penalty to unfair systems, demanding a 57% increase in accuracy to make up for a one-level decline in fairness. This pattern of trade-offs supports the idea that technical performance is not enough to promote adoption: NHS procurement stakeholders consider ethical alignment and interpretability to be crucial.

Taken together, these results contribute to the scholarly discussion of AI in healthcare by offering a holistic framework for approaching AI adoption as a combination of institutional, organizational, and cognitive logics. The thesis specifically advances three areas:

1. *Behavioural Decision-Making in Innovation* – by empirically examining how cognitive biases influence professional adoption of high-risk technologies, this research applies Prospect Theory to the domain of managerial and clinical decision-making.
2. *Strategic and Organizational Innovation Management* – by mapping global patent networks and examining procurement decisions, the thesis clearly shows how innovation ecosystems and institutional mechanisms control the diffusion of technology in the healthcare sector.
3. *Policy and Governance of Emerging Technologies* – the study offers managerial and policy recommendations to support reliable, human-centred innovation by aligning the results with the European Union's AI Act and digital health strategies.

From a managerial standpoint, this work emphasizes that organizational transformation is necessary to accomplish technological transformation in healthcare. Human capital development, open governance practices, and cross-sector cooperation among businesses, organizations, and regulators must be implemented in tandem with investments in digital infrastructures. Businesses in

the medical technology and healthcare industries should implement integrated strategies that combine technological innovation with risk management, ethical compliance, and stakeholder engagement.

From a policy perspective, this research supports that consistent frameworks and standardization procedures remain essential for reducing uncertainty, improving liability, and facilitating responsible AI development and spread. Initiatives aiming to develop a European-wide certification and auditing mechanisms could be a major step towards scaling up trustworthy AI development and implementation within health and care settings.

In conclusion, this doctoral research goes a long way towards harmonizing the two fields of the one bridge – technological innovation on the one side and managerial economics on the other, thus offering a solid, evidence-based ground for comprehending the way risk, trust, and governance interplay to influence the use of AI. Future research could dive deeper into these results to examine longitudinal data on post-adoption outcomes, cross-industry comparisons, and the economic impact of ethical AI application.