D1.1

Final Report - Activity 1: Artificial Intelligence Impact

DOCUMENT	D1.1
PROJECT START DATE	01/10/2021
DURATION	1 YEARS

PROJECT ACRONYM	BRIDGE
PROJECT TITLE	Boosting Remote Interactions and Digital Group Engagement
TOPIC	Artificial Intelligence
DELIVERABLE STATE	V1

DOCUMENT TITLE	**Final Report - Activity 1: Artificial Intelligence Impact**

ABSTRACT SEE EXECUTIVE SUMMARY

Table of Contents

Executive Summary

This deliverable presents a comprehensive analysis of artificial intelligence from legal and ethical perspectives within the European Union framework, examining the evolving regulatory landscape and its implications for the deployment of AI across critical societal sectors. The study adopts a multidisciplinary approach, combining legal document analysis with AI-supported research methodologies to assess the current state of AI governance in Europe and its broader implications for technological development and social welfare.

The analysis reveals that the European Union has positioned itself as a global leader in AI regulation through the introduction of categorising AI systems into four distinct risk levels: unacceptable, high risk, limited risk, and minimal risk. The regulatory framework emphasises the protection of fundamental rights, safety standards, and ethical principles, distinguishing the European approach from more innovation-focused regulatory strategies adopted in other jurisdictions such as the United States and China. EU AI Act[1], the world's first comprehensive legal framework governing artificial intelligence. This landmark legislation establishes a risk-based classification system.

The theoretical framework analysis demonstrates the complexity of defining artificial intelligence and its operational boundaries, highlighting the limitations of current AI systems despite their remarkable capabilities in specific domains. The taxonomic classification of AI systems reveals diverse applications ranging from Artificial Narrow Intelligence in contemporary implementations to the theoretical prospects of Artificial General Intelligence, each presenting distinct regulatory and ethical challenges.

An examination of existing legal instruments reveals the interconnected nature of AI regulation with established frameworks, particularly the General Data Protection Regulation. Gaps remain in addressing the unique characteristics of algorithmic decision-making systems. Member state variations in implementation strategies reflect diverse national priorities and institutional capacities, while international comparative analysis reveals divergent regulatory philosophies that may impact global AI development trajectories.GDPR[2]), product liability directives, and intellectual property laws. These instruments provide foundational principles that extend into AI governance, although

The ethical dimensions analysis identifies four critical areas of concern: autonomy and human agency, transparency and explainability, privacy and data protection, and accountability and liability frameworks. These ethical considerations are particularly acute given the "black box" nature of many advanced AI systems and the increasing delegation of consequential decisions to algorithmic processes.

The sectoral analysis examines AI deployment across five critical domains—healthcare, finance, public administration, security, and energy—revealing both transformative potential and significant ethical challenges. In healthcare, AI applications range from diagnostic assistance to personalised medicine, but they also raise concerns about patient privacy, algorithmic bias, and clinical accountability. Financial sector applications, such as algorithmic trading and credit assessment, pose challenges related to discrimination, systemic risk, and transparency. The deployment of AI in public administration raises concerns about democratic accountability and procedural fairness. At the same time, security applications raise fundamental questions about surveillance, civil liberties, and the

delegation of lethal decision-making to artificial systems. Energy sector implementations focus on optimisation and sustainability but create concerns about justice, privacy, and systemic vulnerability.

1 Introduction

Artificial Intelligence (AI) represents one of the most rapidly evolving areas of technological innovation, characterised by significant advancements in developing increasingly sophisticated models capable of learning, reasoning, and decision-making[1]. The accelerating pace of AI development is driven by both academic and industrial efforts to achieve higher levels of automation and intelligence across various domains. However, this race to advance AI capabilities also surfaces several pressing issues that must be addressed.

Among the most significant challenges associated with the widespread development and deployment of AI are energy consumption, data governance, and the inadequacy of current legal and regulatory frameworks. Firstly, the computational demands of training and operating large-scale AI systems, particularly those based on deep learning, have resulted in high energy requirements[2]. This raises concerns about the environmental impact of AI technologies and the need for sustainable practices in their development and use.

Secondly, AI systems depend heavily on vast quantities of data, often derived from individuals and organisations. This dependency creates complex ethical and legal challenges related to data ownership, consent, privacy, and the potential misuse of sensitive information[3]. Ensuring transparency, accountability, and fairness in data collection and processing becomes essential in this context.

Thirdly, existing laws and regulations frequently fall short in addressing the unique characteristics and implications of AI. Legal systems struggle to define responsibility and liability for AI-generated decisions, protect individuals from algorithmic discrimination, and provide clear rules for emerging AI applications[4]. As AI becomes more embedded in critical sectors such as healthcare, finance, and public administration, the urgency to develop robust regulatory responses increases.

On a global scale, different regions have adopted diverse approaches to regulating AI. Some countries prioritise innovation and economic competitiveness, often with minimal regulatory constraints, while others—particularly within Europe—emphasise ethical governance and the protection of fundamental rights. This divergence reflects broader societal and cultural differences, highlighting the importance of analysing AI development within specific legal and political contexts. This deliverable focuses on the European Union's approach, recognising its emphasis on ethical AI development, strong data protection standards, and efforts to create a unified regulatory framework.

1 https://www.jbs.cam.ac.uk/2025/ai-is-changing-innovation-heres-how/
2https://www.technologyreview.com/2025/05/20/1116327/ai-energy-usage-climate-footprint-big-tech/
3 https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem
4 https://www.ibm.com/think/topics/algorithmic-bias

1.1 Scope and Objectives of the Deliverable

The primary objective of this deliverable is to offer a detailed and structured overview of the current state of ethical and legal considerations in the field of Artificial Intelligence within the European Union. As of the time of writing, the deliverable aims to consolidate information on relevant legislative initiatives, ethical guidelines, and institutional efforts that contribute to shaping AI governance in Europe.

This analysis will serve multiple purposes. It aims to support policymakers, researchers, and practitioners in understanding the evolving landscape of AI regulation. Additionally, it will provide insights into the main trends and challenges currently influencing AI development from an ethical and legal perspective. The deliverable does not propose normative solutions, but instead seeks to describe and critically assess the current framework and the trajectory it outlines.

Moreover, by situating the discussion in the European context, the deliverable highlights how different member states interpret and implement common principles. It also considers the broader implications of EU policies on international AI development, given the EU's influential role in setting global standards, particularly in areas such as privacy and data protection.

1.2 Methodology

The methodology adopted for this deliverable comprises a dual approach that combines legal document analysis with AI-supported research tools. The first step involved a thorough review of existing legislative texts, policy documents, and official publications at both the EU level and within selected member states, including the Artificial Intelligence Act[5], the General Data Protection Regulation (GDPR)[6], national AI strategies, and other relevant instruments.GDPR[2])

The use of advanced AI tools for deep research and document analysis complemented this legal and regulatory review. These tools facilitated the extraction of thematic trends, the identification of recurrent legal issues, and the mapping of ethical concerns reflected in public discourse and institutional documentation. The integration of AI tools into the research process enabled the analysis of a broader and more diverse range of sources, thereby enhancing the comprehensiveness and accuracy of the findings.

5 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024.

6 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), OJ L 119, 4.5.2016, p 1.

In addition to normative texts, this methodology also involved consulting academic literature, reports from expert groups, and public consultations related to AI governance. This triangulation of sources ensures a more comprehensive understanding of the current situation, while also reflecting on the potential and limitations of using AI to analyse its own legal and ethical environment.

1.3 Context of AI Development in the European Framework

The European Union presents a particularly complex and nuanced environment for AI development due to its multi-level governance structure, diverse legal systems, and strong commitment to fundamental rights. Unlike other jurisdictions that may prioritise technological growth and market-driven innovation, the EU has positioned itself as a global leader in promoting trustworthy, ethical, and human-centric AI.

This orientation is reflected in several policy and legislative initiatives. The European Commission has proposed a regulatory framework designed to address risks associated with specific AI applications through a risk-based approach. The proposed AI Act[7] distinguishes between unacceptable, high-risk, and limited-risk AI systems, establishing obligations that vary depending on the level of risk posed to individuals and society.

Furthermore, the EU's approach is heavily influenced by existing data protection laws, particularly those which set a high standard for the lawful processing of personal data. The emphasis on transparency, accountability, and individual rights extends to AI regulation, presenting both opportunities and constraints for developers and organisations.GDPR[2]

Complicating the landscape further is the need to ensure harmonisation among member states, each of which may have differing capacities, priorities, and interpretations of ethical norms. National AI strategies vary in scope and ambition, and although the EU seeks to coordinate these efforts, discrepancies persist in their implementation.

In addition to legal and policy considerations, Europe's historical, cultural, and political values shape its stance on AI. There is a strong emphasis on protecting democratic values, preventing discrimination, and ensuring social inclusion. These elements contribute to making the European context particularly demanding for AI development, yet also potentially more sustainable and equitable in the long term. In the following Table, we report a summary of the European AI Development context.

Aspect	Description

7 https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

Governance Structure	Multi-level governance involving both EU institutions and individual member states, requiring harmonisation across diverse legal systems.
Policy Orientation	Emphasis on ethical, trustworthy, and human-centric AI over rapid, market-driven technological innovation.
Key Legislative Initiative	The AI Act proposal applies a risk-based classification: unacceptable, high-risk, and limited-risk AI systems, with corresponding legal obligations.
Data Protection Influence	Strongly influenced by the GDPR, ensuring high standards for lawful data processing, transparency, and individual rights.
Member State Disparities	National AI strategies vary in their ambition and implementation, resulting in challenges in achieving uniform application across the EU.
Societal Values	Democratic principles, anti-discrimination, and social inclusion are central to Europe's regulatory stance.
Development Implications	Developers face a more complex regulatory environment, but with potential for more sustainable and socially responsible outcomes.

Table 1 – Summary of AI Development Context within the EU

2 Theoretical Framework

2.1 Defining Artificial Intelligence: Technical and Conceptual Boundaries

Artificial intelligence, at its core, comprises a set of technologies that empower computers to execute a diverse array of advanced functions, including the ability to interpret visual information, comprehend and translate spoken and written language, analyse vast datasets, and generate recommendations[8]. These capabilities often rely on sophisticated machine learning algorithms and intricate neural networks that enable systems to learn from data, discern patterns, and perform tasks without being explicitly programmed for every single scenario. The operational principle behind much of AI involves simulating cognitive functions associated with the human mind, such as the capacity for perception, logical reasoning, continuous learning, effective problem-solving, and even creative endeavours.

However, despite the remarkable advancements in the field, current AI systems are bound by several fundamental limitations. These include a lack of deep understanding of the world and common-sense reasoning, hindering their ability to operate beyond the patterns learned from data. True creativity and originality remain elusive, as machines struggle to innovate or envision abstract concepts beyond their training data. The absence of inherent ethical frameworks and moral reasoning poses a significant challenge, as AI decisions are based on learned patterns that can inadvertently perpetuate biases present in the training data. Furthermore, the interpretability of some AI models remains limited, functioning as "black boxes" that make it difficult to understand the reasoning behind their conclusions, which is particularly critical in sensitive domains like healthcare and law. The effectiveness of AI is also heavily dependent on the quality and quantity of training data, as biased or incomplete datasets can lead to skewed or inaccurate outputs. The training of sophisticated AI models demands substantial computational power and energy, raising environmental concerns and limiting accessibility[9]. Moreover, while AI excels in specific domains, transferring knowledge to unrelated tasks remains challenging, and systems are vulnerable to adversarial attacks that can mislead their output. Finally, current AI lacks genuine emotional intelligence and empathy, and exhibits limited ability to learn and adapt in real-time to dynamic environments.

The very definition of "intelligence" and "artificiality" in the context of these technologies remains a subject of ongoing debate and exploration. The technical boundaries of AI are also shaped by factors such as processing capacity, which continues to evolve and define the frontiers of what AI can achieve. Various taxonomies categorise AI capabilities, including algorithmic intelligence, cognitive robotics, and autonomous agents, reflecting the diverse ways in which AI is being developed and utilised. The concept of a socio-technical boundary further emphasises that the impact of AI extends beyond mere technology, requiring interdisciplinary thinking to understand its effects on society.

8 https://www.ibm.com/think/topics/artificial-intelligence
9 https://news.mit.edu/2025/explained-generative-ai-environmental-impact-0117

Ultimately, it is crucial to recognise that AI is a product of human creation, rooted in mathematics and code, and not an inherently mystical or uncontrollable force.

2.2 Taxonomies of AI Systems and Their Applications

The diverse nature of AI systems necessitates various methods of classification to understand their capabilities and applications better. One common taxonomy categorizes AI based on its level of intelligence and abilities, distinguishing between Artificial Narrow Intelligence (ANI), which focuses on specific tasks; Artificial General Intelligence (AGI), which aims for human-level intelligence across a wide range of functions; and Artificial Super Intelligence (ASI), a hypothetical AI that would surpass human intelligence.

Another useful taxonomy classifies AI by its functionality and operational characteristics. Reactive machines represent the most basic level, responding to specific inputs without memory of the past. Limited memory machines possess a temporary understanding of past events, allowing for more complex interactions. Theory of mind machines, currently theoretical, would understand that other entities have beliefs and desires. Finally, self-aware machines, also theoretical, would possess consciousness and an understanding of themselves.

AI systems can also be categorised by their technical learning methods. Supervised learning involves training algorithms on labelled datasets to make predictions or classifications. Unsupervised learning focuses on discovering patterns in unlabeled data. Reinforcement learning trains agents to make optimal decisions by learning from feedback in the form of rewards or penalties.

From a functional perspective, AI algorithms can be grouped by their primary capabilities, including prediction, classification, association, and optimisation, which represent the core tasks that AI is employed to perform. The NIST AI Use Taxonomy provides a human-centred approach, categorising AI based on its contributions to human-AI tasks, including content creation, decision-making, detection, and digital assistance.

The categorising AI systems based on their potential risk levels: unacceptable risk (prohibited), high risk (subject to stringent requirements), limited risk (with transparency obligations), and minimal risk (largely unregulated). General-purpose AI (GPAI) models, which can be used across various applications, are also subject to specific transparency requirements and potential evaluations for systemic risk.EU AI Act[1] introduces a regulatory taxonomy,

Furthermore, AI can be classified hierarchically, starting from the broad field of machine learning, which encompasses deep learning, and more recently, generative AI and large language models (LLMs), such as GPTs. Finally, a taxonomy of trustworthiness for AI categorises systems based on

properties like validity, safety, accountability, explainability, privacy, and fairness, focusing on both the ethical and performance-related aspects of AI.

Illustrative Applications of AI System Taxonomies

The various taxonomies of AI systems are reflected in a wide array of real-world applications across diverse industries and domains. Artificial Narrow Intelligence (ANI) powers many of the AI applications encountered daily. Virtual assistants like Siri, Alexa, and Google Assistant utilise ANI to understand and respond to voice commands. Recommendation engines on platforms such as Netflix and Amazon employ ANI to suggest content based on user preferences. Email services like Gmail utilise ANI for spam filtering, and facial recognition systems rely on ANI for authentication and identification purposes. In transportation, autonomous vehicles utilise ANI for navigation and decision-making, while in healthcare, ANI aids in medical diagnosis. The financial sector leverages ANI for fraud detection, and manufacturing utilises ANI in robotics for automation. Even search engines like Google's RankBrain10 employ ANI to interpret user queries and provide relevant results.

While true Artificial General Intelligence (AGI) remains theoretical, current AI applications offer glimpses into its potential. Future AGI could enable self-driving cars to make complex decisions akin to human drivers, and power customer service systems capable of personalised and empathetic interactions. AGI could also lead to advancements in coding intelligence, allowing systems to generate and improve code, and transform financial services through more accurate modelling and risk assessment. In education, AGI could facilitate truly personalised learning paths, and in exploration, it could power autonomous systems for navigating and understanding complex environments. AI is already having a profound impact on healthcare. In radiology and pathology, AI assists in medical diagnosis by interpreting imaging results. It plays a role in drug discovery by accelerating the identification of potential treatments, and enables personalised treatment plans based on individual patient data. AI-powered chatbots and virtual assistants enhance the patient experience by offering appointment reminders and personalised health advice. In healthcare data management, AI helps to organise and analyse large volumes of patient information. Robotic surgery utilises AI to enhance precision and minimally invasive procedures, and telehealth leverages AI for remote patient monitoring. Furthermore, AI is being explored for its ability to predict disease progression and personalise interventions. The financial industry is also undergoing a significant transformation through the application of AI. AI algorithms are crucial in fraud detection and prevention by analysing transactional data for anomalies. They enhance risk assessment and management by processing vast amounts of financial data, and power algorithmic trading platforms for faster and more efficient execution. AI enables personalised financial services by providing tailored advice and recommendations, and improves the accuracy and speed of credit scoring. In loan processing, AI streamlines tasks such as risk assessment and document verification, and it assists in portfolio management by analysing market conditions. AI is also being utilised for cash flow forecasting to support financial planning. In the realm of transportation, AI is driving innovation across various modes of transportation. Autonomous vehicles, including cars, trucks, and drones, rely on AI for perception, decision-making, and navigation. AI-powered traffic management systems optimise traffic flow and reduce congestion, while route optimisation algorithms enhance efficiency in logistics and delivery services. Predictive maintenance utilises AI to forecast vehicle and infrastructure failures,

10 https://moz.com/learn/seo/google-rankbrain

enabling proactive repairs. AI also plays a role in optimising public transportation systems by analysing passenger flow data and in developing smart parking solutions to improve urban mobility. The education sector is leveraging AI to personalise learning experiences for students through adaptive platforms and intelligent tutoring systems. AI automates administrative tasks for educators, such as grading and scheduling, and facilitates immersive learning experiences through virtual trips. Assistive technologies powered by AI provide support for students with disabilities, and AI tools aid in the creation of smart and engaging content. Manufacturing is undergoing a significant transformation with the integration of AI. Predictive maintenance algorithms analyse sensor data to forecast equipment failures, while AI-enhanced quality control systems improve product inspection and defect detection. AI optimises supply chain management by improving demand forecasting and streamlining logistics. Collaborative robots (cobots) work alongside human workers, enhancing productivity and safety. Generative design technologies use AI to explore and optimise product designs, and digital twins provide virtual replicas of manufacturing processes for real-time monitoring and optimisation. AI also plays a crucial role in demand forecasting, ensuring efficient production planning. In the retail industry, AI is enhancing various aspects of the customer experience and operational efficiency. AI-powered systems optimise inventory management by predicting demand and monitoring stock levels. Demand forecasting algorithms help retailers anticipate customer needs and adjust their forecasts accordingly. Personalised shopping recommendations, driven by AI, enhance customer engagement and drive sales. AI also optimises supply chain and logistics operations to reduce costs and improve efficiency. Chatbots and virtual assistants provide instant customer support and personalised assistance. Visual search capabilities, powered by AI, improve product discovery for online shoppers. Dynamic pricing algorithms enable retailers to adjust prices in real-time based on various factors, while AI-driven systems help detect and prevent fraud, thereby enhancing security. The entertainment industry is leveraging AI to personalise content delivery through recommendation systems on platforms like Netflix, Spotify, and YouTube. AI algorithms are used to generate music, automate video editing processes, and create more realistic non-playable characters (NPCs) in game development. Personalised advertising campaigns are also powered by AI, targeting specific audience segments. Furthermore, AI is used for content analysis to identify trends and understand audience preferences.

3 Legal Framework for AI in Europe

3.1 Overview of the EU AI Act[1]

The European Union has taken a pioneering step in regulating artificial intelligence with the introduction of the The primary objective of this landmark legislation is to foster the development and deployment of trustworthy AI within Europe and beyond, ensuring that AI systems respect fundamental rights, adhere to safety standards, and align with ethical principles. To achieve this, the EU AI Act establishes a risk-based classification system for AI systems, categorising them into four distinct levels: unacceptable risk, high-risk, limited risk, and minimal or no risk.EU AI Act1, the world's first comprehensive legal framework governing AI.

AI practices deemed to pose an unacceptable risk, such as government-run social scoring systems, AI that manipulates human behaviour to cause significant harm, and the untargeted scraping of facial images to create facial recognition databases, are strictly prohibited under the Act. AI systems identified as high-risk, which include those used in critical infrastructure, education, employment, law enforcement, and other sensitive areas, are subject to stringent requirements before they can be placed on the market or put into service. These requirements encompass the implementation of risk management systems, adherence to data governance measures, ensuring record-keeping and transparency, providing for human oversight, and ensuring accuracy, robustness, and cybersecurity. AI systems classified as limited risk, such as chatbots and deepfake detection tools, are subject to specific transparency obligations, requiring providers to ensure that users are aware they are interacting with an AI system. AI systems categorised as minimal or no risk, such as AI-enabled spam filters and video games, face minimal to no regulatory obligations under the Act.

The the provision of technical documentation, information for downstream providers, and summaries of copyrighted data used for training. High-impact GPAI models that may pose systemic risks are subject to more thorough evaluations and reporting obligations. To oversee the implementation and enforcement of the EU AI Act, the European Union has established an EU AI Office. The Act features a staggered applicability timeline, with prohibitions on unacceptable-risk AI systems taking effect in February 2025, and the majority of provisions becoming fully applicable by August 2026. Non-compliance with the EU AI Act can result in significant financial penalties, ranging up to 7% of a company's global annual turnover, underscoring the EU's commitment to ensuring adherence to the new regulations11.EU AI Act1 also addresses general-purpose AI (GPAI) models, which are subject to transparency requirements, including

11 https://www.quinnemanuel.com/the-firm/publications/initial-prohibitions-under-eu-ai-act-take-effect/

3.2 Existing Legal Instruments and Their Application to AI

Beyond the European legal instruments are relevant to AI technologies. The General Data Protection Regulation (minimisation, transparency, individual rights (access, explanation, erasure), and data protection by design are highly relevant to the development and deployment of AI systems that utilise personal data. Data protection authorities across Europe, such as the CNIL in France, are actively issuing recommendations and taking enforcement actions to ensure GDPR compliance in the context of AI. Product liability laws are also increasingly being applied to AI technologies, particularly with the recent updates to the EU Product Liability Directive. This directive now explicitly includes software and AI-integrated products within its scope, addressing liability for defective products, including those with insufficient software updates or cybersecurity vulnerabilities. The concept of "defect" is being adapted to consider the self-learning capabilities of AI systems, raising essential considerations for manufacturers and providers of AI-powered products. Furthermore, intellectual property laws, including copyright and patents, are being examined in the context of AI-generated content and the use of copyrighted material for AI training. The question of who owns the intellectual property rights in AI-generated works and whether the use of copyrighted data for training AI models constitutes infringement is a subject of ongoing legal debate and litigation across Europe and internationally. Courts are grappling with defining authorship and inventorship in the age of AI, and new legal frameworks may be necessary to address these evolving challenges.EU AI Act1, several existing GDPR2) plays a crucial role in governing the processing of personal data used in AI systems. While the

3.3 Member State Variations in AI Regulation

While the harmonised approach to AI regulation across member states, variations exist in their specific postures and implementation strategies. EU member states are in the process of designating the necessary authorities for AI regulation, including market surveillance authorities, notifying authorities, and national public authorities. The structural and design choices for these authorities can vary, as seen in Spain's centralised Spanish Artificial Intelligence Supervisory Agency (AESIA) compared to Finland's decentralised model, which utilises existing market surveillance bodies.EU AI Act1 aims for a

Many EU member states have also developed their own national AI strategies, even before the full enforcement of the the ethical development and deployment of AI, enhancing AI education and skills, and investing in AI infrastructure. For instance, Germany's national AI strategy aims to establish "AI

made in Germany" as a global trademark for cutting-edge and secure AI applications12, while France's strategy, under the "France 2030" plan13, seeks to bolster the nation's AI capabilities and drive economic success. Italy's AI strategy for 2024-2026 focuses on supporting the creation and adoption of AI applications, promoting research, and enhancing the contextual conditions for generating AI value14. Many of these national strategies align with the broader EU objectives of promoting "AI made in Europe" and achieving strategic autonomy in this critical technological domain. Even before the full implementation of the EU AI Act, some national data protection authorities, such as the Dutch DPA, have begun monitoring and supervising AI algorithms, particularly concerning transparency and fairness. The rapid advancements in AI, particularly in areas such as generative AI, are also prompting member states to revise their national strategies to adapt to these new developments.EU AI Act1. These strategies often focus on areas such as fostering AI research and innovation, promoting

3.4 Comparative Analysis with International Jurisdictions

The regulatory landscape for AI varies significantly across the globe. The United States has adopted a more sector-specific approach to AI regulation, emphasising innovation and relying on non-binding guidelines such as the AI Bill of Rights15 and the NIST AI Risk Management Framework16. This decentralised framework contrasts with the EU's comprehensive and legally binding AI Act. China, on the other hand, employs a combination of top-down government control and decentralised innovation in its AI regulation. While also adopting a risk-based approach, China's regulations have a strong focus on national interests, including specific rules for generative AI, deep synthesis technologies, and algorithmic recommendations, as well as requirements for security reviews and algorithm filing17. Canada's proposed Artificial Intelligence and Data Act18 (AIDA) reflects a risk-based approach similar to that of the EU, but it has faced legislative delays and potential changes. The United Kingdom

12 https://www.bmbf.de/EN/Research/EmergingTechnologies/ArtificialIntelligence/artificialintelligence_node.html

13 https://dig.watch/resource/frances-ai-national-strategy

14 https://www.agid.gov.it/en/agenzia/stampa-e-comunicazione/notizie/2024/07/22/italian-strategy-artificial-intelligence-2024-2026

15 https://www.ibm.com/think/topics/ai-bill-of-rights

16 https://www.nist.gov/itl/ai-risk-management-framework

17 https://law.asia/china-ai-regulations-legislation-compliance-future-prospects/

18 https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document

initially adopted a pro-innovation, non-binding framework for AI regulation19, relying on existing sectoral regulators. Still, there are indications of a possible shift towards more structured regulation. Despite the diverse approaches, a common thread across many international AI regulatory frameworks is the prioritisation of mitigating harm, ensuring transparency, and establishing accountability in AI systems.

19 https://www.gov.uk/government/publications/ai-regulation-a-pro-innovation-approach

4 Ethical Dimensions of AI Implementation

4.1 Autonomy and Human Agency

The integration of human and artificial intelligence solutions presents a future that is both brimming with opportunities and challenges across various fields, including healthcare, manufacturing, and decision-making processes. This collaboration has the potential to enhance productivity, foster innovation through combined problem-solving, and enable more informed and efficient real-time decision-making in complex and dynamic environments. In industrial settings, the synergy of human expertise and machine intelligence enables optimised workflows, reduced errors, and enhanced worker safety by delegating hazardous tasks to autonomous systems.

However, the path to seamless and ethical human-AI integration is fraught with potential problems. One significant ethical consideration is the potential for job displacement as AI technologies become more capable of automating routine and even complex tasks20. Privacy concerns also arise with the increasing collection and analysis of vast amounts of personal data by AI systems. Algorithmic bias, stemming from biased training data, poses a risk of perpetuating and amplifying societal inequalities in areas such as hiring, lending, and even the criminal justice system. Ensuring adequate human oversight and control over increasingly autonomous AI systems is another critical ethical challenge. The evolving roles of humans and AI in collaborative environments will require individuals to adapt their skills and focus on higher-level problem-solving and creative tasks, necessitating continuous learning and upskilling initiatives. Furthermore, challenges in communication and coordination between humans and AI systems, as well as the need to develop mutual understanding and trust between human and machine teammates, must be addressed for successful collaboration to occur.

4.2 Transparency and Explainability

Transparency and explainability have emerged as fundamental ethical principles in the context of artificial intelligence, playing a crucial role in building trust, ensuring accountability, mitigating bias, and enabling informed decision-making. Transparency in AI refers to the openness and clarity regarding how AI systems work, encompassing aspects such as explainability, governance, and accountability. Explainability, on the other hand, focuses on the ability of AI systems to provide understandable reasons for their decisions or actions. Together, these principles are crucial for ensuring that AI systems operate ethically and responsibly, enabling users and stakeholders to understand the rationale behind AI-driven outcomes, particularly in high-stakes environments.

However, achieving transparency and explainability in complex AI models presents significant challenges. Many advanced AI models, particularly those based on deep learning, function as "black boxes," making their decision-making processes difficult to interpret. There can also be a trade-off

20 https://www.nexford.edu/insights/how-will-ai-affect-jobs

between model complexity, which often enhances performance, and the goal of making these models understandable. Furthermore, intellectual property concerns may limit the disclosure of proprietary information about AI model architectures and training data.

Despite these challenges, various ethical frameworks and initiatives are promoting the development of explainable AI. The such as the emphasise transparency and explainability. The SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) are being developed to help interpret complex models, even if they don't fully reveal the internal reasoning. These efforts aim to bridge the gap between complex AI models and human understanding, fostering greater trust and responsible AI deployment.EU AI Act1 itself mandates stringent transparency and explainability requirements for high-risk AI systems, requiring providers to ensure their operation is comprehensible to users. Techniques like SHAP (GDPR2 includes a "right to explanation" for automated decisions, and frameworks OECD Principles4 on AI and the NIST AI Risk Management Framework also

4.3 Privacy and Data Protection Implications

The increasing reliance of AI systems on vast amounts of data, including personal information, has profound ethical implications for privacy and data protection. The collection, storage, and use of personal data by AI systems raises concerns about unauthorised data use, the handling of sensitive biometric data, covert data collection practices, and the potential for algorithmic bias to lead to privacy violations. Striking a balance between the utility of data for AI innovation and the imperative to protect individual privacy is a key ethical challenge.

To address these concerns, ethical guidelines for data use in AI have been proposed, emphasising the importance of obtaining informed consent, ensuring transparency about data processing practices, utilising anonymisation techniques to protect privacy, employing representative and unbiased sampling methods for training data, adhering to relevant data protection regulations, and maintaining high data quality. Bias in AI data, stemming from societal biases reflected in training datasets, poses a significant ethical risk to both privacy and fairness, potentially leading to discriminatory outcomes and the perpetuation of inequalities. Robust data protection and security measures, including data minimisation, purpose limitation, and privacy-enhancing technologies, are crucial for mitigating these risks and ensuring the ethical handling of personal data in AI systems.

4.4 Accountability and Liability Frameworks

Determining accountability and establishing liability frameworks for actions and errors of AI systems present complex ethical and legal challenges, particularly in the context of "black box" algorithms and increasingly autonomous systems. The traditional models of accountability, often relying on clear lines of human control and intent, struggle to adapt to AI systems that operate with varying degrees of autonomy based on algorithmic decision-making. Establishing clear responsibility for the outcomes of AI systems, especially in critical applications like healthcare and autonomous vehicles, is essential for ensuring ethical behaviour and building trust.

The challenge of attributing liability becomes even more pronounced when considering the "black box" nature of many advanced AI models, where the decision-making process is opaque and difficult to understand. Responsibility for AI actions could potentially lie with various stakeholders, including AI users, their managers, the companies employing AI, AI developers, and AI vendors, depending on the specific context and the nature of the error or harm caused. The evolving legal landscape is grappling with these issues, with discussions around adapting existing product liability and tort law principles to the unique characteristics of AI systems. The EU has been particularly active in this area, with the recently revised Product Liability Directive explicitly including software and AI-integrated products. A proposed AI Liability Directive aims to ease the burden of proof for victims by creating a rebuttable presumption of causality. Establishing clear accountability and liability frameworks is crucial not only for providing recourse to individuals harmed by AI errors but also for fostering a culture of responsible AI development and deployment.

5 AI in Critical Sectors

The integration of artificial intelligence (AI) technologies across critical sectors represents one of the most significant technological transformations of the contemporary era. These sectors—healthcare, finance, public administration, security, and energy—constitute the foundational infrastructure of modern society, where algorithmic decision-making processes increasingly influence outcomes that directly impact human welfare, economic stability, and societal governance. The deployment of AI systems within these domains necessitates rigorous examination of both the transformative potential and the inherent ethical challenges that emerge from the intersection of advanced computational capabilities and critical societal functions.

This chapter provides a systematic analysis of AI implementation across five critical sectors, examining the specific applications, benefits, and ethical considerations that characterise each domain. The study emphasises the paramount importance of establishing robust ethical frameworks that can adequately address the complex moral dimensions arising from the deployment of AI in contexts where algorithmic decisions carry significant consequences for individuals and society as a whole.

5.1 AI for Healthcare

5.1.1 Applications and Implementation

The healthcare sector has witnessed extensive adoption of AI technologies across multiple domains, fundamentally altering diagnostic procedures, treatment protocols, and patient care delivery mechanisms. Machine learning algorithms, particularly deep learning architectures, have demonstrated remarkable efficacy in medical imaging analysis, enabling automated detection of pathological conditions in radiological examinations, histopathological specimens, and ophthalmological assessments. Computer-aided diagnosis (CAD) systems now assist clinicians in identifying malignancies, cardiovascular abnormalities, and neurological disorders with accuracy levels that often surpass those of human experts.

Natural language processing (NLP) technologies have revolutionised clinical documentation and decision support systems, facilitating automated extraction of clinical insights from electronic health records (EHRs), medical literature, and patient-generated data. Predictive analytics models enable early identification of patient deterioration, adverse drug reactions, and hospital readmission risks, thereby supporting proactive interventions and resource optimisation.

Personalised medicine represents another significant domain of AI application, where algorithmic analysis of genomic data, biomarkers, and phenotypic characteristics enables tailored therapeutic interventions. AI-driven drug discovery platforms accelerate the identification of novel therapeutic compounds and optimise clinical trial design through patient stratification and endpoint prediction.

5.1.2 Ethical Considerations

The implementation of AI in healthcare raises profound ethical concerns that demand careful consideration. **Patient privacy and data protection** constitute primary ethical imperatives, as AI systems require access to extensive personal health information for training and operation. The aggregation and analysis of sensitive medical data create vulnerabilities to privacy breaches and unauthorised access, necessitating robust data governance frameworks and encryption protocols.

Algorithmic bias poses a significant concern, particularly about health disparities across demographic groups. AI systems trained on datasets that inadequately represent diverse populations may perpetuate or exacerbate existing healthcare inequities, leading to suboptimal outcomes for marginalised communities. The lack of diversity in training data can result in diagnostic algorithms that perform poorly for specific ethnic groups, gender identities, or socioeconomic strata.

Clinical accountability and liability issues emerge when AI systems contribute to medical decision-making processes. The question of responsibility for AI-assisted diagnoses or treatment recommendations presents complex legal and ethical challenges, particularly in cases where algorithmic outputs conflict with clinical judgment or result in adverse outcomes.

Informed consent becomes increasingly complex when AI systems are involved in patient care, as patients may lack understanding of algorithmic processes and their implications. The **transparency and explainability** of AI decision-making processes are crucial for maintaining patient trust and facilitating informed consent. Yet, many advanced AI systems operate as "black boxes" with limited interpretability.

5.2AI for Finance

5.2.1 Applications and Implementation

The financial sector has been at the forefront of AI adoption, leveraging algorithmic capabilities to enhance operational efficiency, risk management, and customer experience. **Algorithmic trading** systems utilise machine learning models to execute high-frequency transactions, optimise portfolio management, and identify market opportunities through pattern recognition and predictive analytics. These systems process vast quantities of market data in real-time, enabling rapid responses to market fluctuations and arbitrage opportunities.

Credit risk assessment has undergone a fundamental transformation through the AI-enabled analysis of traditional and alternative data sources. Machine learning algorithms evaluate creditworthiness by analysing patterns in financial behaviour, social media activity, and transactional data, thereby enabling more accurate risk scoring and expanded access to credit for previously underserved populations.

Fraud detection and prevention systems utilise anomaly detection algorithms and behavioural analytics to identify suspicious transactions and account activities. These systems continually learn from transaction patterns to adapt to evolving fraud schemes, thereby reducing false positive rates while maintaining security effectiveness.

Robo-advisors and automated financial planning platforms democratise access to investment advice by leveraging AI-driven portfolio optimisation and personalised financial recommendations. Natural language processing enables sophisticated chatbots and virtual assistants that provide customer support and financial guidance.

5.2.2 Ethical Considerations

The implementation of AI in finance presents significant ethical challenges that intersect with economic justice, privacy, and systemic stability. **Algorithmic discrimination** in lending and insurance decisions represents a paramount concern, as AI systems may perpetuate historical biases embedded in training data, leading to unfair denial of financial services based on protected characteristics or proxy variables.

Financial inclusion and exclusion dynamics are profoundly influenced by AI systems that determine access to credit, insurance, and investment opportunities. While AI can potentially expand financial access through alternative data analysis, it may also create new forms of exclusion for individuals who lack digital footprints or whose behavioural patterns deviate from algorithmic expectations.

Market manipulation and systemic risk concerns arise from the widespread deployment of algorithmic trading systems, which may contribute to market volatility, flash crashes, or coordinated market abuse. The interconnectedness of AI-driven trading systems creates a potential for systemic risks that could destabilise financial markets.

Privacy and data exploitation issues are particularly acute in financial AI applications, where systems aggregate and analyse comprehensive profiles of individual economic behaviour. The commodification of personal financial data raises questions about consent, data ownership, and the appropriate limits of behavioural surveillance.

Transparency and accountability in financial AI decisions are crucial for regulatory compliance and consumer protection; however, many AI systems lack sufficient explainability to enable meaningful oversight and dispute resolution.

5.3AI for Public Administration

5.3.1 Applications and Implementation

Public administration is increasingly relying on AI technologies to enhance service delivery, optimise resource allocation, and improve governance processes. Automated decision-making systems streamline benefit eligibility determinations, permit approvals, and regulatory compliance assessments, thereby reducing processing times and administrative costs while potentially improving consistency in decision outcomes.

Predictive policing and public safety applications utilise machine learning algorithms to forecast crime patterns, optimise patrol deployment, and identify individuals at risk of recidivism. These systems analyse historical crime data, demographic information, and environmental factors to inform law enforcement's resource allocation and intervention strategies.

Smart city initiatives integrate AI technologies across urban infrastructure, including traffic management systems, energy distribution networks, and public transportation optimisation. Machine learning algorithms process sensor data and citizen interactions to enhance urban planning and improve service delivery efficiency.

Digital government services utilise AI-powered chatbots, document processing systems, and citizen engagement platforms to enhance the accessibility and responsiveness of public services. Natural language processing enables the automated analysis of citizen feedback and sentiment monitoring across digital platforms.

5.3.2 Ethical Considerations

The deployment of AI in public administration raises fundamental questions about democratic governance, citizen rights, and the legitimate scope of governmental algorithmic power. **Procedural fairness and due process** concerns arise when AI systems make decisions that affect citizen access to benefits, services, or legal status. The opacity of algorithmic decision-making can undermine citizens' ability to understand, challenge, or appeal automated decisions.

Democratic accountability becomes problematic when public officials delegate decision-making authority to AI systems without maintaining adequate oversight or a thorough understanding of algorithmic processes. The erosion of human discretion in public administration may compromise the responsiveness and contextual sensitivity that characterise effective governance.

The surveillance and civil liberties implications are particularly significant in the public safety applications of AI, where predictive policing systems may disproportionately target certain communities and perpetuate discriminatory enforcement patterns. The aggregation of citizen data across government systems creates a potential for comprehensive surveillance that may chill free expression and association.

Digital divide and access equity concerns arise when AI-driven government services disproportionately benefit digitally literate citizens, potentially excluding vulnerable populations who lack technical skills or digital access. The automation of government services must consider the differential impacts on diverse citizen populations.

Transparency and public trust requirements necessitate that citizens comprehend how AI systems impact government decisions that affect their lives. The legitimacy of democratic governance depends on citizens' ability to understand and evaluate the algorithmic processes that shape the implementation of public policy.

5.4AI for Security

5.4.1 Applications and Implementation

AI technologies have become integral to contemporary security frameworks, encompassing national security, cybersecurity, and physical security domains. **Threat detection and intelligence analysis** systems utilise machine learning algorithms to process vast quantities of signals intelligence, identify patterns indicative of security threats, and predict potential attack vectors. These systems analyse communication metadata, behavioural patterns, and network traffic to detect anomalous activities that may indicate malicious intent.

Autonomous weapons systems represent an emerging and controversial application of AI in military contexts, where machine learning algorithms enable weapon platforms to identify, track, and engage targets with minimal human intervention. These systems raise profound questions about the delegation of lethal decision-making to artificial systems.

Cybersecurity applications utilise AI for real-time threat detection, automated incident response, and vulnerability assessment. Machine learning models analyse network traffic patterns, endpoint behaviours, and code repositories to identify malware, intrusion attempts, and system vulnerabilities before they can be exploited.

Border security and surveillance systems integrate AI-powered facial recognition, biometric analysis, and behavioural detection technologies to monitor and control population movements. These systems process video feeds, identification documents, and biometric data to identify individuals of interest and detect suspicious activities.

5.4.2 Ethical Considerations

The deployment of AI in security contexts generates particularly acute ethical dilemmas due to the intersection of technological capabilities with fundamental rights and liberties. **Lethal autonomous weapons systems** raise profound moral questions about the permissibility of delegating life-and-death decisions to artificial systems. The development of weapons that can select and engage targets without

meaningful human control challenges established principles of international humanitarian law and raises concerns about accountability for unlawful killings.

Mass surveillance and erosion of privacy represent significant concerns as AI systems enable unprecedented scales of population monitoring and behavioural analysis. The aggregation of surveillance data across multiple systems creates comprehensive profiles of individual activities and associations, potentially chilling democratic participation and dissent.

Racial and demographic profiling concerns arise when AI security systems exhibit biased performance across different population groups. Facial recognition systems have demonstrated higher error rates for women and individuals with darker skin tones, leading to disproportionate targeting and false identifications that can result in wrongful detention or harassment.

Human rights and civil liberties implications extend beyond privacy to encompass freedom of movement, association, and expression. AI-enabled security systems may facilitate authoritarian control mechanisms that suppress political opposition and minority rights.

International stability and arms race dynamics emerge as nations compete to develop advanced AI-enabled security capabilities, potentially destabilising existing strategic balances and creating pressures for accelerated deployment of insufficiently tested systems.

5.5 AI for Energy

5.5.1 Applications and Implementation

The energy sector has adopted AI technologies to optimise grid operations, enhance the integration of renewable energy, and improve energy efficiency across consumption patterns. Smart grid management systems utilise machine learning algorithms to balance electricity supply and demand in real-time, predict equipment failures, and optimise energy distribution across complex network topologies. These systems process data from smart meters, sensors, and weather forecasts to enable dynamic pricing and load balancing.

Renewable energy optimisation applications utilise AI to enhance the efficiency of solar and wind energy installations through predictive maintenance, performance optimisation, and integration planning. Machine learning models forecast renewable energy generation based on weather patterns and optimise energy storage systems to mitigate supply variability.

Energy consumption analytics enable both utilities and consumers to identify efficiency opportunities through pattern recognition and predictive modelling. AI systems analyse building

energy usage, industrial processes, and transportation patterns to recommend optimisation strategies and automate energy management decisions.

Oil and gas exploration applications utilise AI for seismic data analysis, drilling optimisation, and reservoir modelling, enabling more efficient resource extraction while potentially reducing environmental impacts through improved targeting and reduced waste.

5.5.2 Ethical Considerations

The implementation of AI in the energy sector raises complex ethical questions related to environmental sustainability, energy justice, and technological dependency. **Environmental impact and climate change** considerations require careful evaluation of whether AI-enabled energy systems contribute to or mitigate environmental degradation. While AI can optimize renewable energy systems and reduce consumption, the computational requirements of AI systems themselves consume significant energy and may offset environmental benefits.

Energy justice and accessibility concerns arise when AI-optimised energy systems create differential access to affordable and reliable energy services. Dynamic pricing algorithms may disproportionately burden low-income consumers who lack flexibility to adjust consumption patterns, while innovative grid technologies may be deployed primarily in affluent areas.

Privacy and behavioural surveillance issues arise as smart meter systems and energy management platforms collect detailed data about household activities and occupancy patterns. The granular analysis of energy consumption data can reveal intimate details about personal behaviours and lifestyle choices.

Economic displacement concerns affect energy sector workers whose roles may be automated through the implementation of AI. The transition to AI-enabled energy systems must consider the social and economic impacts on communities that depend on traditional energy industries.

Systemic vulnerability and resilience questions arise as energy infrastructure becomes increasingly dependent on AI systems that may be vulnerable to cyberattacks, algorithmic failures, or adversarial manipulation. The concentration of energy management functions in AI systems creates potential single points of failure that could compromise energy security and stability.

6 Conclusion

This comprehensive analysis of artificial intelligence from legal and ethical perspectives within the European Union framework reveals the profound complexity of governing transformative technologies that increasingly permeate critical societal infrastructure. The research demonstrates that the European Union's approach to AI regulation represents a paradigmatic shift toward prioritising ethical considerations, fundamental rights protection, and social responsibility over purely innovation-driven technological development.

The . At the same time, its emphasis on transparency, accountability, and human oversight sets essential precedents for global AI governance. However, the analysis also reveals significant implementation challenges, including the need for harmonisation across member states with varying institutional capacities and regulatory philosophies.EU AI Act1 emerges as a landmark achievement in technology governance, establishing the world's first comprehensive regulatory framework for artificial intelligence. Its risk-based approach provides a nuanced method for addressing the diverse applications and potential harms associated with AI systems

The theoretical and taxonomic analysis of AI systems underscores the importance of maintaining definitional clarity and technical precision in regulatory frameworks, as the boundaries and capabilities of artificial intelligence continue to evolve rapidly. The limitations of current AI systems, particularly in terms of explainability and interpretability, present ongoing challenges for regulatory compliance and ethical implementation across all analysed sectors.

The intersection of AI governance with existing legal instruments demonstrates both the adaptability and limitations of traditional regulatory approaches when confronted with novel technological capabilities. While frameworks such as the critical foundational principles for AI governance, the unique characteristics of algorithmic decision-making systems require specialised regulatory responses that address emerging risks and ethical challenges.GDPR2 provide

The ethical dimensions identified throughout this analysis—autonomy, transparency, privacy, and accountability—represent fundamental considerations that transcend sectoral boundaries and technological specifications. These principles must be embedded as primary design constraints rather than secondary considerations in AI development and deployment processes, ensuring that technological advancement serves to enhance rather than diminish human dignity and social welfare.

The sectoral analysis reveals that the deployment of AI across critical domains generates both sector-specific and cross-cutting ethical challenges. Healthcare applications demonstrate the potential for AI to enhance diagnostic accuracy and treatment personalisation while raising fundamental questions about patient autonomy and clinical responsibility. Financial sector implementations illustrate the tension between efficiency gains and concerns about algorithmic discrimination and systemic risk. Public administration applications underscore the crucial importance of upholding democratic accountability and procedural fairness in automated decision-making systems. Security sector deployments present existential questions about the delegation of consequential decisions, particularly those involving the use of force, to artificial systems. Energy sector applications demonstrate the complex relationship between technological optimisation and considerations of justice, sustainability, and resilience.

The comparative international analysis reveals divergent regulatory philosophies that reflect broader cultural and political values regarding the appropriate relationship between technological innovation and social governance. The European emphasis on precaution, rights protection, and ethical consideration contrasts with more permissive approaches adopted elsewhere, potentially creating competitive tensions while also establishing critical normative standards for global AI development.

Looking forward, several critical challenges emerge from this analysis. First, the rapid pace of AI technological development, particularly in areas such as generative artificial intelligence and large language models, creates ongoing pressure to adapt regulatory frameworks and ethical guidelines to address the novel capabilities and risks they present. Second, the global nature of AI development and deployment necessitates increased international coordination to prevent regulatory arbitrage and ensure consistent protection of fundamental rights across jurisdictions. Third, implementing and enforcing comprehensive AI regulations requires significant institutional capacity building and technical expertise, which may vary considerably across member states and sectors.

The research also highlights the importance of maintaining adaptive and responsive regulatory approaches that can evolve with technological advancement while preserving core ethical principles and rights protections. This requires ongoing dialogue between technologists, policymakers, civil society organisations, and affected communities to ensure that regulatory frameworks remain relevant and practical.

Furthermore, the analysis demonstrates the critical importance of investing in AI literacy and public understanding of algorithmic systems, as democratic governance of AI technologies requires informed public participation in policy-making processes. The complexity of AI systems and their societal implications necessitate educational initiatives that enable citizens to engage meaningfully with questions about the appropriate role of artificial intelligence in society.

The European Union's approach to AI governance, as analysed in this deliverable, offers a valuable model for striking a balance between innovation and ethical considerations and rights protection. However, the success of this approach will ultimately depend on effective implementation, continued adaptation to technological advancements, and a sustained commitment to the human-centric values that underpin the regulatory framework.

In conclusion, the governance of artificial intelligence represents one of the defining challenges of the contemporary era, requiring sustained interdisciplinary collaboration and ethical reflection to ensure that these powerful technologies serve to enhance human welfare and social justice. The European framework provides essential foundations for this endeavour. Still, continued vigilance, adaptation, and commitment to fundamental values will be necessary for navigating the complex landscape of AI governance in the years ahead. The stakes of these decisions extend beyond technological optimisation to encompass the foundational structures and values that define democratic society and human dignity in an age of algorithmic mediation.

7 References

1. European Commission. (2021). Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).

2. European Union. (2016). General Data Protection Regulation (EU) 2016/679.

3. High-Level Expert Group on AI. (2019). Ethics Guidelines for Trustworthy AI.

4. OECD. (2019). OECD Principles on Artificial Intelligence.

5. The White House Office of Science and Technology Policy. (2022). Blueprint for an AI Bill of Rights.

6. United Nations. Reports on Lethal Autonomous Weapons Systems (LAWS).

7. International Energy Agency. (2017). Digitalisation and Energy.