FRIEDRICH NAUMANN FOUNDATION For Freedom.
Central Europe

ZAVOD 14
zavod za sožitje in napredek

# Understanding the Socioeconomic and Environmental Impact of Artificial Intelligence: **A Comprehensive Policy Roadmap for the European Union**

Dejan Ravšelj • Eva Murko • Matej Babšek • Nina Tomaževič • Aleksander Aristovnik

# Understanding the Socioeconomic and Environmental Impact of Artificial Intelligence: A Comprehensive Policy Roadmap for the European Union

**Authors**
Dejan Ravšelj, Eva Murko, Matej Babšek, Nina Tomaževič, Aleksander Aristovnik

**Editor**
Zavod 14

**Publisher**
Zavod 14, zavod za sožitje in napredek

**Graphic Design by**
spelabecaj.si, freepik.com

**Supported by**
Friedrich-Naumann-Stiftung für die Freiheit
Truman Haus,
Karl-Marx-Straße 2,
14482 Potsdam-Babelsberg,
Germany

/freiheit.org
/FriedrichNaumannStiftungFreiheit
@FNFreiheit

**Contact**
Phone: +49 30 22 01 26 34
Fax: +49 30 69 08 81 02
email: service@freiheit.org

**Date**
November 2024

ZAVOD 14
zavod za sožitje in napredek

Zavod 14, zavod za sožitje in napredek
Cankarjeva 4 • 1000 Ljubljana

/zavod14
@stirinajst

Zavod 14, zavod za sožitje in napredek, is a non-profit (ELF full member) organization and has its headquarters in Celje (Slovenia). Zavod 14 promotes social liberal ideas (balancing between individual liberty and social justice) and protect liberal values (eg. democracy, the rule of law, social development, good governance etc.) Zavod 14`s mission is strengthening the civil society initiative, integration and cooperation of the interested public, transmission of perspectives of interested stakeholders to state and other institutions, cooperation in the preparation and implementation of politics, and contribution in joining the civil society initiative into international integrations.

The publication is supported by the Friedrich Naumann Foundation for Freedom. The Friedrich Naumann Foundation for Freedom is not responsible for the content of this publication, or for any use that may be made of it. The views expressed herein are those of the author(s) of the publication alone. These views do not necessarily reflect those of the Friedrich Naumann Foundation for Freedom.

FRIEDRICH NAUMANN FOUNDATION For Freedom.
Central Europe

# Contents

## Dejan RAVŠELJ

Dejan Ravšelj, PhD, is an Assistant Professor at the Chair of Economics and Public Sector Management, Faculty of Public Administration, University of Ljubljana, Slovenia. His research interests primarily focus on public sector economics, public sector management, and performance measurement. Recently, he has been actively involved in research on generative artificial intelligence, such as ChatGPT, and its implications for higher education and the future labour market. He is regularly involved as a leader or member of research teams in various domestic and international projects covering diverse research areas. He publishes professional and scientific articles and presents them at a range of domestic and international professional and scientific conferences.

## Eva MURKO

Eva Murko is a Young Researcher and doctoral candidate at the Faculty of Public Administration, University of Ljubljana. Her research interests intertwine public governance reforms and models, the digital transformation of public administration, and the use of artificial intelligence in public administration, with particular attention to its relationship with public governance and socio-economic welfare. She is involved in national research projects, collaborates with international bilateral partners, and frequently participates in international research conferences.

## Matej BABŠEK

Matej Babšek, PhD, is a Research Assistant at the Chair of Economics and Public Sector Management, Faculty of Public Administration, University of Ljubljana, Slovenia. His research interests encompass interdisciplinary studies of public administration, including e-Government with a focus on artificial intelligence, administrative reforms, and the legal and organisational aspects of the work of public institutions. He is a member of various project teams. He publishes professional and academic articles and presents them at a range of domestic and international professional and academic conferences.

## Nina TOMAŽEVIČ

Nina Tomaževič, PhD, is an Associate Professor at the Faculty of Public Administration, University of Ljubljana, Slovenia. She teaches Theory of Organization, Human Resources Management, Management in Public Sector, Organizational and Managerial Development, and Sustainable Development. Her latest research interests encompass organizational aspects of AI introduction in public institutions as well as topics on sustainability. She is an author of numerous scientific and professional articles, monographs and parts of monographs and has actively participated in international and domestic conferences. She is a member of different project teams and professional and scientific associations.

## Aleksander ARISTOVNIK

Aleksander Aristovnik, PhD, is the executive director of Zavod 14 and a Full Professor at the Faculty of Public Administration, University of Ljubljana, Slovenia. His research focuses on public sector economics and management, disruptive technologies, public finance, macroeconomics, and EU economics. He has participated in over seventy international conferences and led or collaborated on numerous research projects funded by the Agency for Research and Innovation of Slovenia (ARIS), including "Developing an AI Readiness Framework for Citizen-Centered Public Governance." He has published extensively in domestic and international journals and is a member of organizations such as the European Economic Association (EEA) and International Network for Economic Research (INFER). He has also served on the Slovenian Officials' Council.

# 1 Introduction

*The evolution of artificial intelligence (AI) from a theoretical concept to a transformative force has dramatically reshaped industries, economies, and societies. Its capacity to enhance productivity, improve decision-making, and optimise processes has led to widespread adoption across various sectors. AI's origins date back to the 1950s. However, despite the initial excitement, early AI developments encountered significant limitations due to the technological constraints of the time, leading to periods of stagnation known as the "AI winters." It was not until the 2000s, with the advent of more powerful computing capabilities, big data, and the refinement of machine learning techniques, that AI truly began to advance at an unprecedented pace. AI now encompasses a wide range of subfields, including machine learning (ML) and deep learning (DL), which have made it possible for AI systems to recognise patterns, make decisions, and even learn from experience. These systems are increasingly capable of adapting to complex and unpredictable environments, opening new possibilities across multiple domains. From healthcare and education to agriculture, manufacturing, and the public sector, AI has emerged as a powerful tool for solving intricate problems and driving efficiency in ways that were previously unimaginable. In the European Union (EU), the integration of AI has seen significant progress, particularly in countries like Germany, the Netherlands, and Italy, which lead the way in adoption and innovation. Smaller nations such as Estonia are also making remarkable strides, demonstrating that AI's potential is not limited to the largest economies.*

*This growing reliance on AI underscores its role in shaping the future of various sectors, but it also presents a series of challenges that require careful consideration. As AI becomes more embedded in both the public and private sectors, understanding its ethical, legal, and sustainable implications becomes crucial. The EU, with its commitment to technological leadership, ethical standards, and sustainability, is playing a central role in addressing these challenges. Its approach to AI regulation, as exemplified by the EU AI Act, aims to balance innovation with safety and ethical oversight. At the same time, emerging trends such as the potential development of Artificial General Intelligence (AGI) and Artificial Super Intelligence (ASI) pose new questions about the long-term societal impacts of AI and its ability to operate beyond human control. These advancements, while offering solutions to some of the world's most complex challenges, also necessitate a thoughtful and deliberate approach. Policymakers and developers must guide AI's development to align with human values, emphasising safety, transparency, and societal benefit.*

*Hence, the goal of this publication is to provide a holistic analysis of AI's multifaceted impacts on European society, the economy, and the environment. As AI increasingly integrates into public* *and private life, it is critical to ensure that its deployment aligns with the EU's core values of human dignity, inclusivity, and sustainability. The publication highlights the immense potential of AI to drive innovation and competitiveness within the EU but also emphasises the need to mitigate risks related to workforce displacement, social inequality, privacy concerns, and the environmental costs associated with AI technologies. This report is divided into three key chapters, each exploring the distinct yet interconnected implications of AI—societal, economic, and environmental—providing a comprehensive overview of both the potential and the challenges that this powerful technology brings.*

*The chapter on societal implications examines how AI is revolutionising core areas of public life, such as healthcare and education. It highlights AI's ability to optimise services, improve quality of life, and create more equitable opportunities through personalised education and healthcare solutions. However, the chapter also addresses significant risks, such as the perpetuation of biases, threats to personal privacy, and the widening digital divide. It emphasises the importance of governance frameworks like the EU's AI Act to ensure that AI technologies are deployed in a fair, transparent, and accountable manner. By exploring AI's dual role in enhancing societal growth and posing ethical dilemmas, the chapter provides a balanced view of its broader social impact.*

*Secondly, the chapter on economic implications delves into how AI is reshaping labour markets, driving productivity, and influencing global competitiveness. It discusses AI's potential to automate routine tasks, displace jobs, and simultaneously create new high-skilled roles, emphasising the need for workforce reskilling and adaptation. AI's ability to boost productivity in sectors like manufacturing, healthcare, and finance offers a pathway to economic growth, but the chapter also highlights disparities in AI adoption between large enterprises and SMEs. Additionally, the chapter explores AI's potential to become a general-purpose technology akin to electricity or the internet, driving continuous innovation. The chapter outlines the specific challenges faced by the EU in realising AI's full economic benefits, including regulatory constraints and investment fragmentation.*

*Lastly, the chapter on environmental implications explores the dual role of AI in both advancing sustainability and posing environmental risks. AI's ability to optimise resource use, improve energy efficiency, and contribute to sectors like agriculture and energy is key to supporting the EU's green transition and climate neutrality goals. However, the chapter also critically assesses the environmental costs of AI, such as the carbon footprint of data centres, the energy-intensive nature of AI model training, and the growing issue of e-waste from AI-related hardware. The concept of "green AI" is introduced as a potential solution, focusing on developing more energy-efficient algorithms and reducing the ecological impact of AI systems. The chapter emphasises the need for global collaboration and regulatory frameworks to align AI development with environmental sustainability goals.*

*This publication provides a comprehensive framework for understanding the far-reaching implications of AI within the EU, outlining a clear roadmap for addressing the complex challenges posed by AI. It offers targeted recommendations to help policymakers harness AI's transformative potential while safeguarding fundamental rights and promoting long-term sustainability across environmental, social, and economic dimensions.*

# 2 The AI Landscape in the EU

## 2.1 Understanding the Fundamentals and Concepts of AI

AI has evolved from a theoretical concept into a transformative technology that is reshaping industries, economies, and societies. Its potential to enhance productivity, improve decision-making, and optimise processes has made AI a powerful tool in addressing some of the most pressing challenges of our time. AI's journey began in the 1950s when Alan Turing introduced the concept of machine intelligence through the "Turing test." The term "artificial intelligence" was later coined by John McCarthy in 1955 (Siebel, 2019). Despite early enthusiasm, progress slowed during the "AI winters" due to insufficient computing power (OECD, 2019). However, the 2000s marked a revival, fuelled by advancements in computing, big data, and machine learning (Siebel, 2019). At its core, AI refers to the capability of machines to perform tasks that typically require human intelligence, such as problem-solving, learning, perception, and decision-making. AI is not a singular entity; it is a broad and complex field that encompasses various subfields, including machine learning (ML) and deep learning (DL), each of which has distinct capabilities and applications. AI systems are designed to pro cess data, learn from it, and adapt their behaviour to achieve specific objectives, often operating in environments that are partially observable or unpredictable.

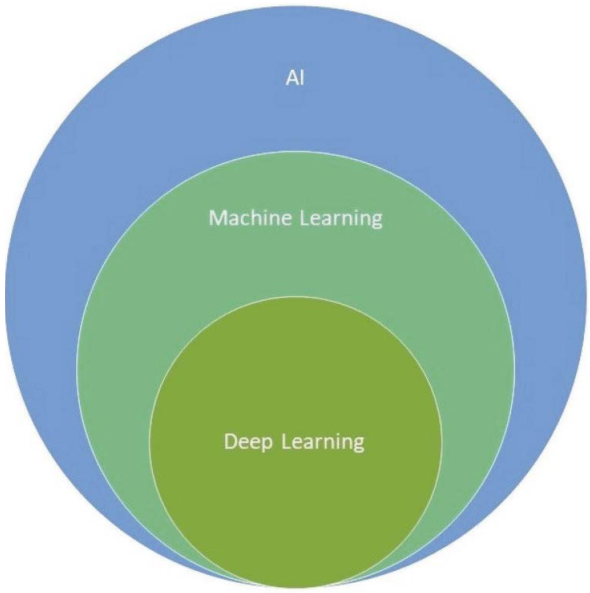
According to the OECD (2024a) latest definition, AI can be described as "a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment." These systems are increasingly being integrated into both public and private sectors, with profound implications for economies, societies, and the environment. However, understanding how AI systems function, how they interact with their societies, and the environment (Aristovnik et al., 2023a), and how they evolve over time is essential for developing policies that promote ethical, sustainable, and responsible AI deployment.

The concept of AI is often broken down into layers or subfields, each with unique attributes that contribute to its overall functioning. The three most commonly referenced layers of AI include (Banfa, 2024):

- **Artificial intelligence (AI):** The broadest term encompassing any machine or system capable of mimicking human cognitive processes. AI systems can be designed to carry out a wide range of tasks, from simple calculations to more complex activities like decision-making, pattern recognition, and language processing.
- **Machine Learning (ML):** A subset of AI that focuses on algorithms that allow machines to learn from data without being explicitly programmed. Machine learning is at the heart of many modern AI applications, enabling systems to identify patterns in data and improve their performance over time.
- **Deep Learning (DL):** A further subset of machine learning, deep learning uses multi-layered neural networks to analyse vast amounts of data. Deep learning has proven especially effective in tasks like image recognition, speech processing, and natural language understanding, where traditional algorithms might struggle.

The relationship between AI, machine learning, and deep learning can be visually represented in Figure 1 to clarify the hierarchical structure of these fields. As shown below, deep learning is a subset of machine learning, which in turn is a subset of AI.

*Figure 1: Relationship between AI, machine learning, and deep learning*



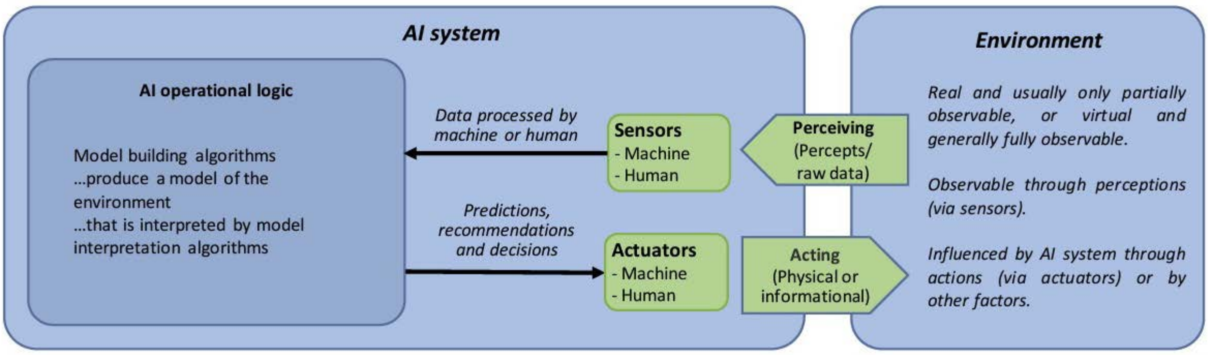

*Source: Banfa, 2024.*

An AI system is designed to interact with its environment and perform tasks based on data inputs and processing mechanisms. Understanding how AI systems function requires a breakdown of the key components that enable them to perceive, process, and act upon data. These components include (OECD, 2019):

- **Sensors:** These may be machine-based (e.g., cameras, microphones) or human inputs that gather raw data from the environment for processing.
- **Perception:** AI systems interpret data collected by sensors to generate an understanding of the environment. This stage involves recognising patterns, identifying objects, or interpreting complex inputs.
- **Actuators:** These are the outputs or actions that AI systems take based on their processing. Actuators may be physical, such as robotic arms, or informational, such as a recommendation generated by a software system.
- **Environment:** AI systems operate within real or virtual environments, which can be either fully or partially observable. The environment influences the AI system's actions, while the AI system also modifies its environment through its decisions and actions.

The following Figure 2 provides a conceptual overview of how an AI system operates, showing how data flows through sensors, is processed, and results in actions that influence the environment. This representation highlights the cyclical relationship between the AI system and its environment. Data is continuously collected and processed to make predictions or recommendations, which lead to actions that further alter the environment, creating a feedback loop that improves the system's functionality over time.

*Figure 2: A high-level conceptual view of an AI system*



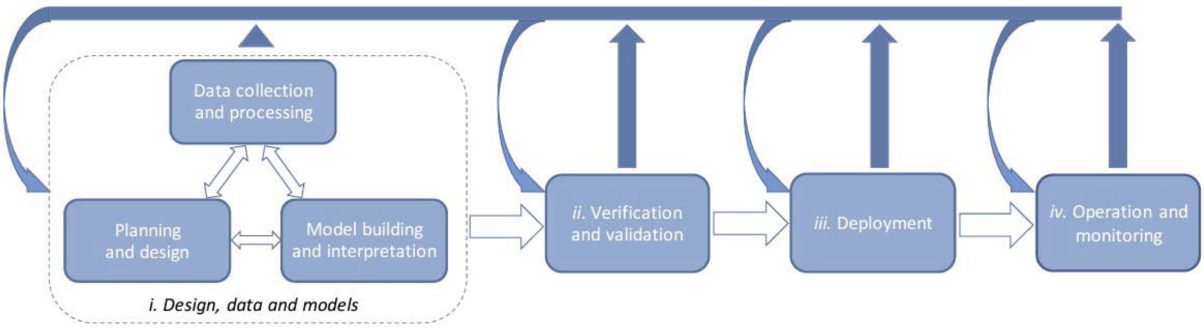

*Source: OECD, 2019.*

An AI system does not function in isolation but follows a structured lifecycle from its initial design to its operation in the real world. Each stage in the AI lifecycle plays a critical role in ensuring that the system performs reliably and responsibly. The AI lifecycle typically consists of the following phases (OECD, 2019):

1. **Design, Data, and Models:** The first stage involves planning and designing the AI system, collecting relevant data, and building models that can interpret the data. These models form the foundation of the AI system's ability to learn and make decisions.
2. **Verification and Validation:** Once the system is designed, it undergoes rigorous testing to ensure it behaves as expected and meets the required performance criteria. Validation ensures the system's outputs align with its objectives, while verification confirms its reliability.
3. **Deployment:** Following successful validation, the AI system is deployed in a real-world environment. At this stage, the system begins interacting with users or other systems, performing the tasks it was designed to accomplish.
4. **Operation and Monitoring:** After deployment, the system's performance is continuously monitored to ensure it remains effective and responsive to changes in its environment. Regular updates or retraining may be required as the system encounters new data or challenges.

This lifecycle is depicted in Figure 3, which illustrates the iterative nature of AI development. By following this structured lifecycle, AI systems are able to evolve and improve, adapting to new environments, data, and tasks. This ensures that AI systems remain functional, efficient, and aligned with their intended objectives over time.

*Figure 3: AI system lifecycle*




*Source: OECD, 2019.*

Understanding the key concepts, operational mechanisms, and lifecycle of AI systems is essential for grasping the full potential of this transformative technology. AI systems operate in dynamic environments, continuously learning and adapting based on the data they process. By recognising how AI systems are designed, deployed, and monitored, we gain valuable insights into how AI can be harnessed to drive innovation, improve decision-making, and address complex societal challenges. Moreover, this knowledge enables policymakers and developers to ensure that AI systems are aligned with ethical standards and societal goals, mitigating risks while maximising benefits. As outlined by the OECD (2019, 2024a), AI systems differ in their levels of autonomy and adaptiveness, making it crucial to design systems that are both robust and transparent. Understanding the lifecycle of AI systems, from design to deployment ensures that these technologies remain functional and adaptable across sectors. With AI influencing global markets and societal structures, a comprehensive understanding of its foundations is vital for fostering responsible development, aligning with sustainability, equity, and competitiveness goals.

## 2.2 European Approach to the Regulation of AI

The establishment of the EU AI Act (Regulation (EU) 2024/1689 of the European Parliament and of the Council, 2024; AI Act) reflects the EU's challenge in balancing the benefits and risks of new technology. A crucial but often overlooked factor in the EU's AI regulation is the role of single market integration, which forms the foundation of the EU's regulatory powers. The market integration paradox highlights the tension between harmonising the market to ensure fair competition and fostering innovation through diversity. This tension is mirrored in AI regulation, where the EU must balance safety and ethical standards with the need to avoid stifling innovation.

At the heart of the digital single market and AI in the EU is its strength in data governance, exemplified by regulations such as the General Data Protection Regulation (Regulation (EU) 2016/679 of the European Parliament and of the Council, 2016; GDPR) and the Data Governance Act (Regulation (EU) 2022/868 of the European Parliament and of the Council, 2022; Data Governance Act). These laws set the standards for data privacy, protection, and sharing, which are crucial for the responsible development of AI. The GDPR, with its focus on data protection principles such as transparency, fairness, data minimisation, and accountability, etc., as shown in Figure 4, also governs automated decision-making and profiling, directly impacting AI applications. These principles ensure that personal data is handled responsibly, which is essential for the ethical use of AI. Additionally, the Digital Services Act (Regulation (EU) 2022/2065 of the European Parliament and of the Council, 2022; Digital Services Act) and the Digital Markets Act (Regulation (EU) 2022/1925 of the European Parliament and of the Council, 2022; Digital Markets Act), while not AI-specific, aim to ensure safer digital spaces and fair, open markets. By regulating digital platforms, these acts address transparency, accountability, and fairness, which are critical for AI-powered services.

*Figure 4: The core principles specified by the GDPR for the processing of personal data*


*Source: OpenMinded, 2022.*

The EU's journey toward establishing the AI Act began with several key initiatives aimed at fostering a harmonised approach to AI development and regulation across Europe. These efforts have laid the groundwork for ethical, trustworthy, and competitive AI in the EU, ensuring that the benefits of AI are maximised while addressing emerging challenges. Below are some of the most important milestones and policy actions that paved the way for the AI Act:

- **European Council Invitation (2017):** The European Council invited the Commission to develop a pan-European approach towards AI (AI Act, 2024), setting the foundation for the EU's AI policy.
- **AI Strategy for Europe (2018):** Published in April 2018, the strategy emphasised the need for a harmonised and collaborative EU-wide approach to maximise AI benefits and address challenges (European Commission, 2018). It highlighted the digital single market as a crucial element for digital growth and innovation.
- **National AI Strategies and the Coordinated Plan (2018):** To prevent fragmented approaches, the Coordinated Plan on AI was launched in 2018 to align Member States' AI strategies, boost investment, and synchronise AI policies across the EU (European Commission, 2021).

- **High-Level Expert Group (HLEG) on AI (2018):** The Commission established the HLEG to draft ethical guidelines for AI. The group published the Ethics Guidelines for Trustworthy AI in 2019 after extensive consultation (European Commission, 2019a).
- **White Paper on AI (2020):** Introduced by the Commission in February 2020, this paper proposed strategies for promoting AI adoption while managing associated risks, laying the groundwork for future AI regulations (European Commission, 2020b).

The AI Act is set to enter into force on 1 August 2024, following a comprehensive development process. Initially proposed by the European Commission on 21 April 2021, the Act has undergone several revisions to address challenges posed by emerging technologies such as general-purpose AI. After extensive discussions and refinements, the final version of the AI Act was published on 26 January 2024, and subsequently adopted by the European Parliament (Sidley, 2024). This regulatory framework aims to ensure that AI systems in the EU are safe, transparent, and uphold fundamental rights while fostering innovation. In Table 1, key aspects of the AI Act are presented.

*Table 1: Key aspects of the AI Act*

| Aspect | Details |
|---|---|
| **Risk-Based Approach** | The AI Act categorises AI systems into risk levels: unacceptable risk, high risk, limited risk, and minimal risk, with regulations tailored to each category. |
| **General Purpose AI (GPAI)** | Special regulations apply to GPAI models, ensuring these widely applicable AI systems comply with safety and ethical standards. |
| **Transparency and Accountability** | AI systems must be transparent, with clear information on how they function and how decisions are made, particularly for high-risk AI. |
| **Ethics and Fundamental Rights** | The Act promotes AI that respects fundamental rights, such as privacy, non-discrimination, and environmental protection, ensuring alignment with the EU's core values. |
| **Compliance and Enforcement** | Companies and organisations deploying high-risk AI systems are required to comply with strict regulations, including risk management, documentation, and reporting. |
| **AI Testing and Certification** | High-risk AI systems must undergo testing, validation, and certification to ensure they meet safety and performance standards before being deployed. |
| **Fines for Non-Compliance** | Non-compliance with the AI Act can result in significant fines, particularly for companies deploying high-risk AI systems without adhering to the necessary regulations. |
| **Support for Innovation** | While ensuring safety, the AI Act encourages innovation by supporting responsible AI development and providing a framework that promotes growth in the AI sector. |

*Source: Own elaboration.*

Figure 5 visually captures the phased implementation of the EU AI Act, highlighting the gradual introduction of key regulations. Starting from the Act's entry into force on 1 August 2024, the timeline shows how restrictions on unacceptable AI systems will take effect 6 months later, followed by the application of codes of conduct after 9 months. The regulatory framework progressively addresses high-risk AI systems, with implementing acts being issued 18 months in, and the full application of the AI Act, including for high-risk AI, after 24 months. By the 36-month mark, all remaining regulations for high-risk AI systems will come into force, ensuring that organisations deploying AI have ample time to meet compliance requirements. This phased approach helps balance the need for oversight with the opportunity for AI innovation.

*Figure 5: Timeline and important deadlines for the AI Act*


*Source: DPO Centre, 2024.*

## 2.3 Recent Advances and Emerging Trends in AI

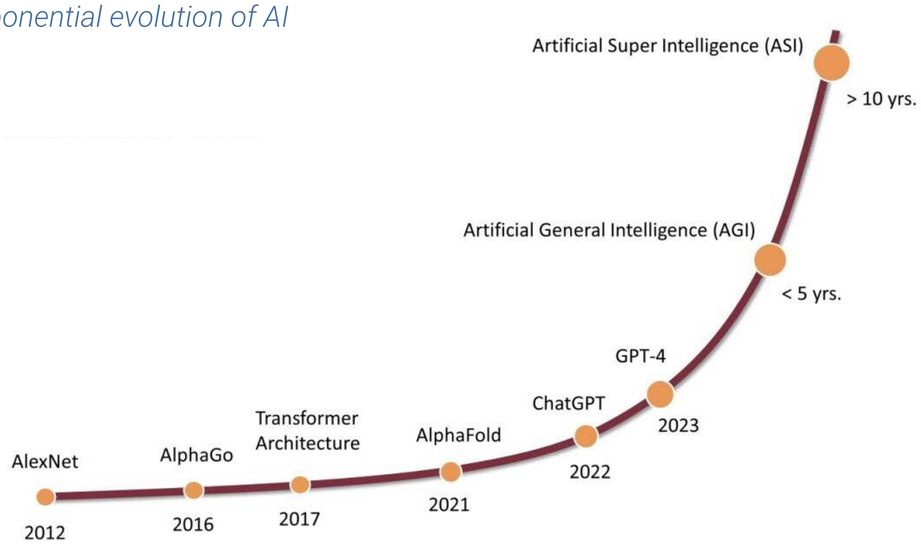
In just over a decade, AI has evolved from a niche interest in academic circles to a powerful force shaping numerous sectors of society. This progression, driven by key breakthroughs, has not only shifted traditional paradigms but also paved the way for future innovations that once seemed like science fiction. A series of pivotal milestones have shaped the development of AI, each one expanding the technology's capabilities and reach (Bhatia, 2024):

- **AlexNet (2012).** The introduction of AlexNet in 2012 marked a breakthrough in AI, particularly in the field of image recognition. By using a deep convolutional neural network (CNN), AlexNet drastically improved the accuracy of image classification, winning the ImageNet competition and outperforming previous models by a significant margin. This milestone proved the effectiveness of deep learning and ushered in an era where AI began to tackle real-world tasks with practical applications. The availability of large datasets and powerful GPUs made it possible to train these deep neural networks, setting the stage for future AI developments across a wide range of industries, including healthcare, automotive, and entertainment.
- **AlphaGo (2016).** In 2016, AlphaGo, developed by DeepMind, made history by defeating a world champion Go player, Lee Sedol, in a five-game series. Go, with its near-infinite possibilities and complexity, was long considered one of the most challenging games for AI to master. AlphaGo's victory was significant because it combined deep neural networks with reinforcement learning and self-play, allowing it to learn strategies through experience rather than being pre-programmed. This accomplishment demonstrated that AI could handle tasks requiring creativity and long-term planning, laying the groundwork for future AI systems capable of tackling complex real-world problems in fields like finance, logistics, and even medical diagnoses.
- **Transformer Architecture (2017).** The introduction of the Transformer architecture in 2017 revolutionized natural language processing (NLP) and set the foundation for a new generation of AI models. The transformer's use of self-attention mechanisms allowed it to process sequences of data more efficiently than previous models, such as recurrent neural networks (RNNs). This architecture became the backbone of powerful models like GPT, BERT, and others, transforming applications in machine translation, text summarization, and conversational AI. The transformer model enabled AI to achieve unprecedented levels of fluency in language understanding and generation, reshaping industries like customer service, education, and content creation.
- **AlphaFold (2021).** AlphaFold's 2021 breakthrough in predicting protein structures represented a monumental leap in the application of AI to biology. Predicting the 3D structure of proteins from their amino acid sequences is a challenge with enormous implications for understanding diseases and developing new drugs. AlphaFold achieved an accuracy level

comparable to experimental methods, solving a problem that had eluded scientists for decades. This success demonstrated the potential of AI not only in computational tasks but also in contributing to scientific discovery, advancing fields such as biochemistry, genetics, and pharmacology by accelerating the pace of research and opening new doors for innovation in medicine.

- **ChatGPT 2022.** In 2022, ChatGPT brought AI-powered conversational agents into everyday use, showcasing the capabilities of large-scale generative models. Based on OpenAI's GPT-3, ChatGPT allowed users to engage in human-like dialogue, making AI accessible in domains such as customer service, education, and creative writing. Its ability to generate coherent, context-aware text responses made it a versatile tool for businesses and individuals alike. ChatGPT's rapid adoption, reaching over 100 million users within months, underscored the growing integration of AI into daily life. This milestone marked a new phase of human-AI interaction, where AI could serve as both an assistant and a collaborator across various tasks.
- **GPT-4. 2023.** The release of GPT-4 in 2023 pushed the boundaries of generative AI even further. As an advanced version of the transformer model, GPT-4 offered improved performance in text generation, translation, and reasoning tasks, thanks to a larger architecture and potential innovations such as the Mixture of Experts (MoE) approach. This model allowed for more efficient processing and task-specific expertise, elevating AI's capabilities in understanding and generating language. GPT -4's versatility in handling complex and creative tasks positioned it as a state-of-the-art tool for industries such as media, research, and software development, setting a new benchmark for what AI could achieve in terms of human-like cognition.

*Figure 6: The exponential evolution of AI*



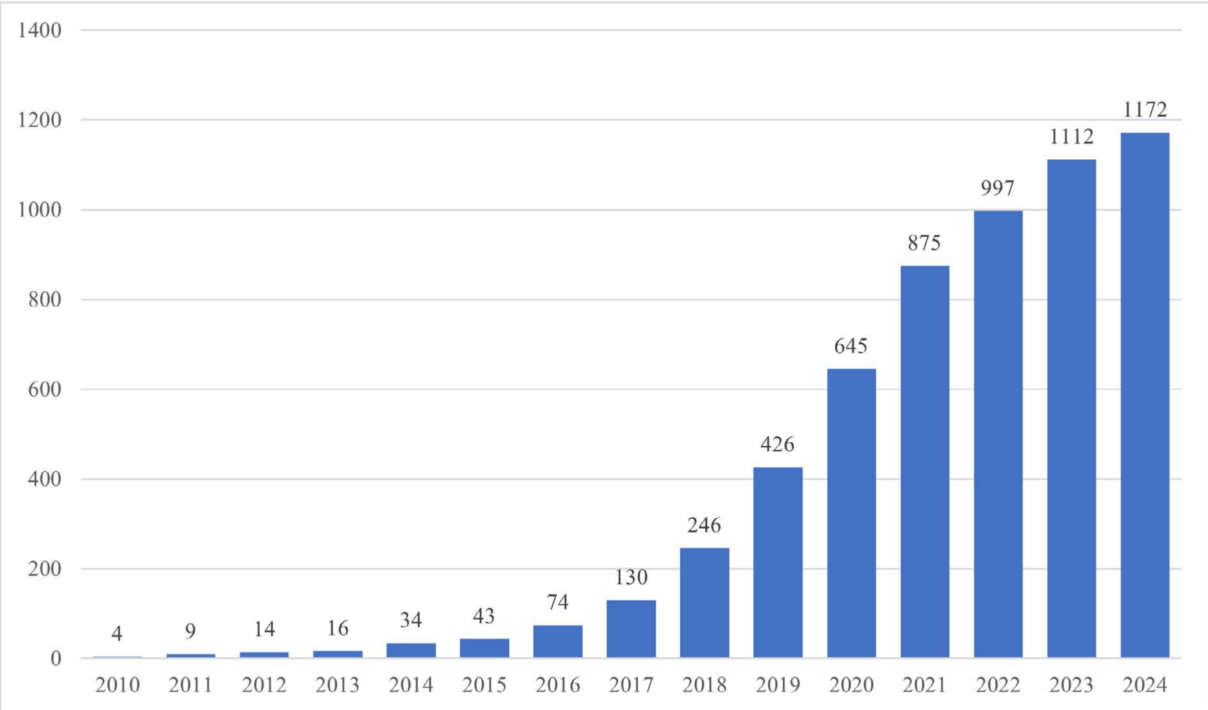

*Source: Bhatia, 2024.*

The future of AI holds exciting prospects, from systems capable of human-like learning and decision-making to advancements that may exceed human intelligence altogether. Experts predict that in the coming years, AI will become more versatile and capable of tackling a broader range of tasks across various domains. However, along with the promise of enhanced productivity and innovation, these advancements raise important ethical and societal questions. As we look to the future, the challenge lies in harnessing AI's potential responsibly while ensuring it aligns with human values and benefits society as a whole (Bhatia, 2024):

- **Artificial General Intelligence (AGI) (< 5 years).** AGI, expected within the next five years, would mark a transformative shift in AI's capabilities, allowing machines to perform tasks across a wide range of domains with human-like learning and reasoning abilities. Unlike current AI systems, which are specialized for narrow tasks, AGI would be able to understand and adapt to new problems without needing extensive retraining. The potential applications of AGI are vast, spanning sectors such as healthcare, education, and engineering, where AI could perform intellectual tasks with the flexibility and creativity of a human being. AGI represents a step closer to true autonomous intelligence, making it possible for AI to function as a general-purpose assistant capable of solving complex, multi-domain challenges.
- **Artificial Super Intelligence (ASI) (> 10 years).** Artificial Super Intelligence (ASI), projected to be achieved in more than 10 years, would surpass human intelligence by an unprecedented margin. ASI could potentially improve its own capabilities recursively, leading to rapid advancements in its cognitive abilities. This intelligence explosion could allow ASI to solve global challenges such as climate change, curing diseases, or developing advanced technologies that are currently beyond human comprehension. However, the development of ASI also raises significant ethical concerns, as the disparity in intelligence between humans and machines could pose risks if ASI's goals are not aligned with human values. Ensuring that ASI development follows strict safety guidelines and ethical frameworks will be crucial to harnessing its potential while mitigating the risks associated with its immense power.

The latest data from the EU reveals a total of 1,172 recorded AI use cases in the public sector, with solutions targeting government-to-government interactions (G2G) (53% of use cases), businesses (G2B) (10% of use cases), and citizens (G2C) (37% of use cases). The adoption of AI in the EU began modestly, with just 4 cases recorded up to 2010 (with the first use case documented in 2003), but has steadily accelerated over the years, particularly from 2017 onwards. From 2010 to 2016, there was only a minimal increase in AI use cases, but starting in 2017, the numbers surged significantly. This sharp rise highlights the expanding role of AI technologies in the public sector, with particularly notable growth after 2017 as yearly increases became much more pronounced (European Commission, 2024b). The year 2017 was a pivotal moment for AI innovation and adoption, driven by the convergence of several key factors. First, the expansion of connected devices, fueled by the rise of sensors and the Internet of Things (IoT), produced vast amounts of data that significantly improved AI systems' analytical capabilities. At the same time, falling computing costs made AI more scalable and accessible, enabling its application across various public sector domains, such as social protection. Additionally, the surge in available data from digitalization, social media, and smartphones provided the essential resources for training AI models, allowing them to function more effectively in public sector contexts. Finally, advancements in machine learning, particularly in deep learning, greatly enhanced AI's ability to recognize patterns, make decisions, and predict outcomes, further strengthening its role in the public sector (Ravšelj & Babšek, 2024; World Economic Forum, 2017). In 2017, the recorded cases jumped to 129 and continued to grow each year, reaching 1,172 by 2024 (Figure 7).

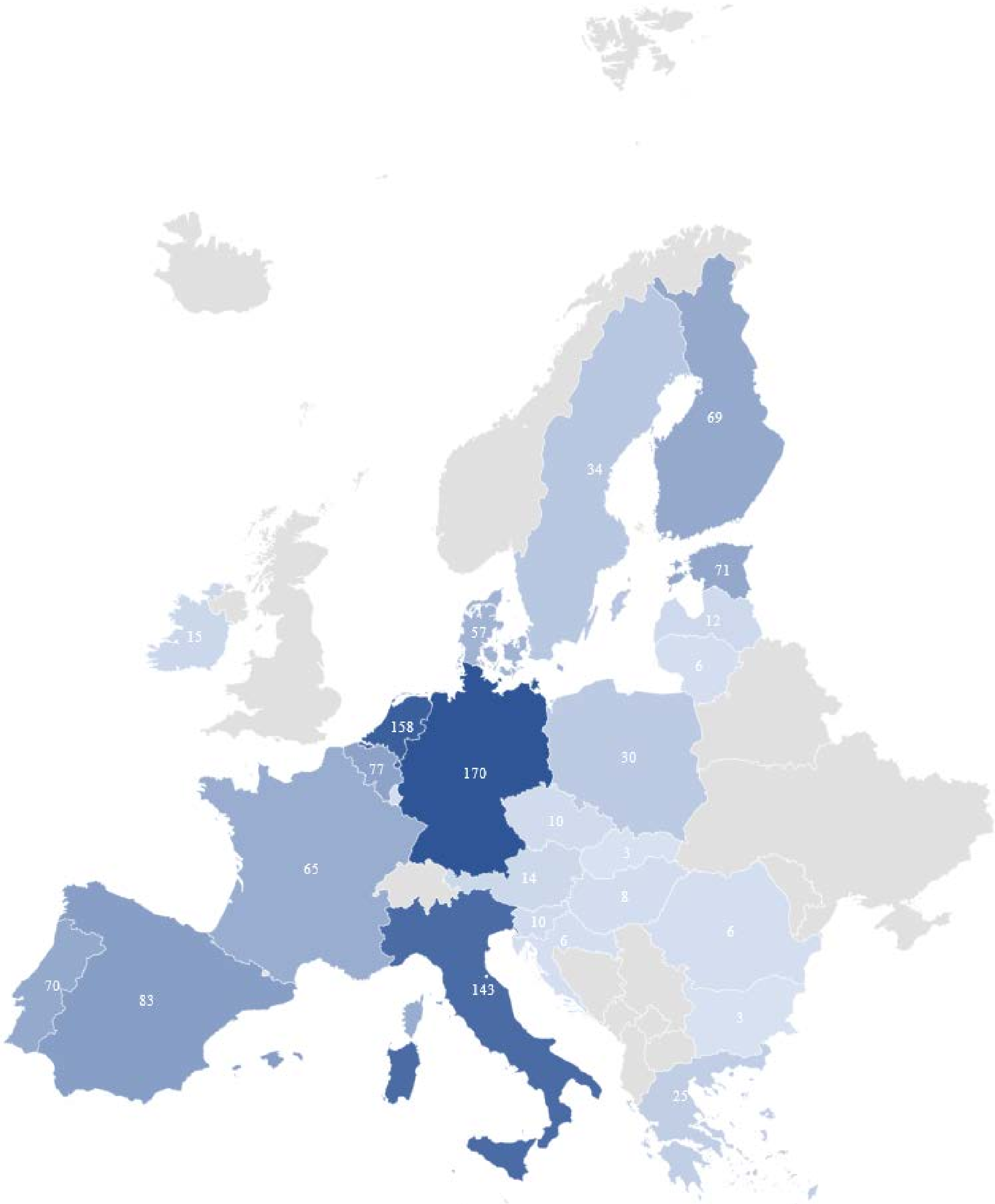
*Figure 7: Cumulative number of AI use cases in the EU*




*Source: European Commission, 2024b.*

The geographical distribution of AI use across EU member states is uneven, revealing substantial differences between countries (Figure 8). Germany leads the list with 170 AI use cases, followed by the Netherlands (158) and Italy (143), indicating that these countries are the most active in AI development and application within the EU. In the mid-range, countries like Spain (83), Belgium (77), and Estonia (71) show moderate activity. Estonia, in particular, stands out as a smaller country with a relatively high number of AI use cases, emphasizing its strong focus on digital transformation. At the lower end of the spectrum, countries like Malta (1), Slovakia (3), and Cyprus (3) report very few AI use cases, which may suggest either a limited focus on AI initiatives or restricted resources for AI development. Eastern European and smaller economies, such as Poland (30), Latvia (12), and Slovenia (10), fall in the lower-middle range, indicating some disparity in AI adoption compared to Western European nations. Overall, this data provides an insightful overview of AI usage distribution across the EU, with larger economies and countries boasting more developed digital infrastructures leading in AI implementation, while smaller or less economically advanced nations contribute less in this regard.

*Figure 8 : Number of AI use cases across the EU member states*




*Source: European Commission, 2024b.*

# 3 Societal Implications of AI

## 3.1 Opportunities and the Need for Ethical Governance of AI

As we transition from the economic implications of AI to its societal impacts, it will become clear that AI's influence extends far beyond financial systems and productivity gains. AI has the potential to reshape the very fabric of society, altering how individuals live, communicate, and participate in public life. AI can drive remarkable advancements in areas such as education, healthcare, and public safety and can be a powerful tool for enhancing such essential services. For instance, by optimising the delivery of healthcare and personalising educational experiences, AI holds the promise of improving the quality of life and creating more equitable opportunities across different communities. It can contribute to smarter, more responsive cities and public systems, helping societies address pressing social issues more efficiently. However, as AI systems become more embedded in everyday life, concerns arise about how they might reinforce societal inequalities. How AI is designed and deployed could perpetuate bias, threaten personal privacy, or exacerbate the digital divide, especially in communities with limited access to advanced technologies. The societal implications of AI demand careful consideration, as the technology can amplify both positive and negative outcomes.

This chapter will explore these dynamics, focusing on both the opportunities AI presents for societal growth, current applications and the risks it poses. We will examine how AI can transform critical areas like healthcare and education, and we will also assess the social risks—such as algorithmic bias and loss of privacy. These risks underscore the need for comprehensive governance frameworks that address the ethical deployment of AI. We will conclude by discussing emerging regulatory approaches, such as the AI Act, and outline key principles for the development of trustworthy and ethical AI systems that prioritise fairness, transparency, and accountability. By examining both sides of the societal impact of AI, this chapter aims to provide a balanced view of how this powerful technology can shape our future.

## 3.2 Harnessing AI for Global Social Progress through Sustainable Development Goals

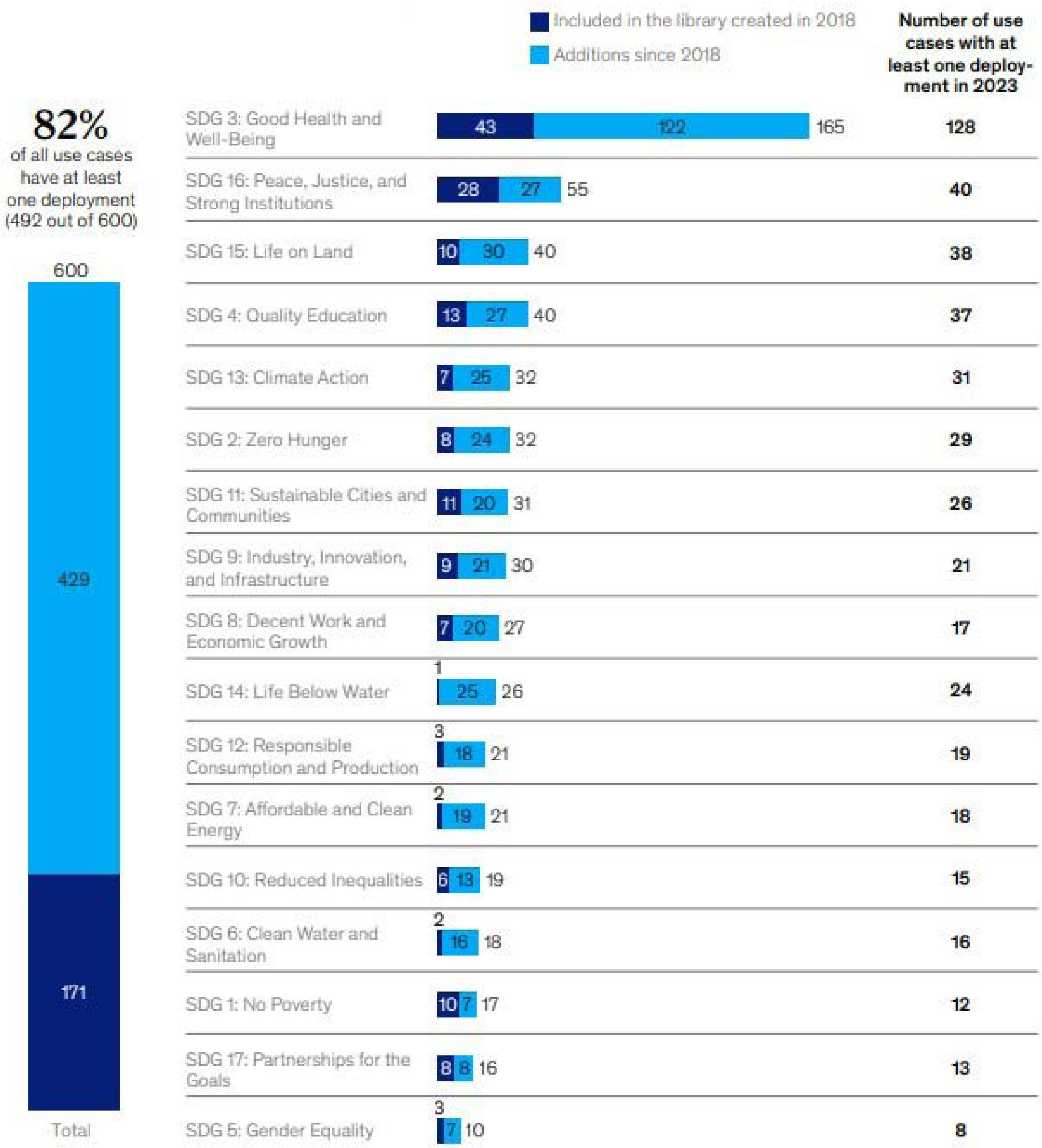
AI is already playing a crucial role in advancing all 17 of the United Nations Sustainable Development Goals (SDGs) (United Nations, n.d.), from poverty eradication to the creation of sustainable cities and ensuring quality education for all. The advent of generative AI has unlocked new opportunities, and looking ahead, there is significant potential for further progress. Emerging tools and platforms are empowering social entrepreneurs, public sector leaders, and private organisations to develop more impactful solutions. However, with this increasing power comes the need to ensure AI is used responsibly, with active management of potential risks to prevent unintended negative consequences (McKinsey Global Institute, 2019; Bankhwal et al., 2024; Murko et al., 2023; Babšek et al., 2024).

AI's global impact extends beyond productivity and economic growth; it is also fostering social good. Current research and development in AI not only validate earlier expectations but suggest the potential for even greater advancements. Significant improvements in AI techniques, particularly generative AI, have expanded the range of challenges AI can address. These advancements include natural language processing, real-time language translation, the synthesis of large datasets, and the creation of diverse content in text, images, and video (Brynjolfsson & Raymond, 2023).

To better understand AI's role in addressing global challenges, we can map innovations and impacts to the SDGs. These goals, consisting of 17 objectives and 169 targets, aim to improve global living conditions and protect the environment. However, the 2023 United Nations report indicates that the world is only on track to achieve 15% of the SDG targets (United Nations, 2023). This shortfall has tangible implications: 2.2 billion people still lack access to safe water and sanitation, 3.3 billion reside in areas highly vulnerable to climate change, and 750 million people are experiencing hunger (Bankhwal et al., 2024). AI is currently being utilised across all SDGs, supporting critical efforts such as protein modelling, drug screening, vaccine development, and optimising the distribution of public services and aid. AI is also addressing supply chain challenges, such as optimising routes for delivering food to remote areas, predicting the long-term effects of climate change, providing early warnings for natural disasters, and equipping frontline aid workers with specialised expertise.

McKinsey Global Institute developed a comprehensive database of AI use cases (Bankhwal et al., 2024), each illustrating a specific challenge that can be addressed using various AI capabilities. In their 2018 report, the database contained around 170 high-potential use cases. By 2024, this number had expanded to roughly 600, reflecting a more than threefold increase. This growth is driven by continued innovation as social impact leaders experiment with AI and as these tools become increasingly accessible and user-friendly. Additionally, the number of real-world AI implementations has grown substantially. While only a small portion of the 170 use cases had been deployed in 2018, by 2024, approximately 490 of the 600 cases, or over 80%, had been put into practice in at least one instance (Bankhwal et al., 2024) (Figure 9).

*Figure 9: Number of AI-enabled use cases identified per UN SDG*




*Source: Bankhwal et al., 2024.*

According to the report by McKinsey Global Institute (Bankhwal et al., 2024), some SDGs have more widely recognised potential, such as Good Health and Well-Being (SDG 3) and Quality Education (SDG 4). Below are the potential and existing deployments for SDG 3 and SDG 4.

### Current Initiatives Related to SDG 3: Good Health and Well-Being

SDG 3 focuses on enhancing global well-being and ensuring healthy lives for all. Its key targets include lowering maternal mortality rates, combating infectious diseases such as HIV/AIDS, tuberculosis, and malaria, and providing universal access to sexual and reproductive health services, family planning, and education (United Nations, n.d.).

AI is now deeply integrated into various medical research and healthcare processes. Prominent AI applications in this field include protein modelling, genome sequencing, CT scan analysis, visual aids, and vaccine development. The healthcare sector is particularly suited for AI integration compared to other SDGs due to its advanced technological infrastructure, the relative abundance of data, and the ease of measuring health outcomes. Despite this, there remain significant opportunities to use AI to address under-prioritized areas, such as treating neglected communicable diseases and preventing substance abuse (Bankhwal et al., 2024).

AI is revolutionising healthcare by enhancing the ability of computers to learn, reason, and apply logic, which opens up transformative possibilities for diagnosis and treatment. Through collaboration among scientists, clinicians, engineers, and researchers, AI systems are being developed that deliver reliable, accurate, and efficient healthcare solutions (Tai, 2020). There are numerous areas showcasing AI's positive impact on the healthcare sector (Figure 10).

*Figure 10: AI use cases in healthcare*


*Source: Shuliak, 2024.*

There are notable advancements in (1) ***diagnostics and treatment,*** where AI enables faster and more precise diagnostic processes. For instance, systems like IBM's Watson have demonstrated the ability to analyse vast amounts of medical data swiftly, providing diagnostic suggestions and potential treatment options. This technology can assess digital results from physical exams and offer comprehensive evaluations of possible conditions, significantly enhancing the speed and accuracy of medical decision-making. Additionally, IBM's Watson for Oncology analyses patient data to recommend personalised treatment plans, improving decision-making by integrating knowledge from leading cancer centres. AI efficiently (2) ***processes large datasets,*** as seen with Google's DeepMind, which helps clinicians identify at-risk patients. In UK mammography cases, their AI reduced false positives by 25% without missing true diagnoses. AI is also improving the quality of life for vulnerable populations through (3) ***therapeutic robots*** designed to offer social support. These AI-powered robots, often used to assist the elderly and physically challenged, provide more than just practical help with daily tasks. They can also alleviate loneliness, reduce anxiety, and lower blood pressure, improving overall well-being. By serving as companions, these robots contribute to enhanced social interaction and emotional health for seniors. In addition, AI contributes to one of the major challenges in healthcare: (4) ***reducing human error.*** Fatigue and emotional distractions can impair healthcare professionals' performance, sometimes leading to serious consequences. AI technologies are not prone to such limitations, delivering consistent, reliable outcomes and minimising the risk of error. By automating and streamlining certain processes, AI helps improve efficiency and accuracy across the healthcare system. Furthermore, (5) ***AI-assisted surgical technologies*** are raising the bar for precision. Tools like the da Vinci surgical system enable healthcare professionals to perform minimally invasive surgeries with remarkable accuracy, reducing trauma, blood loss, and recovery time for patients. While still under the guidance of surgeons, these systems enhance procedural outcomes and alleviate patient anxiety by ensuring a less invasive approach. (6) ***Radiology*** is another field where AI has made significant strides. From the early days of CT and MRI scanning, AI has continued to evolve diagnostic imaging techniques. The development of advanced algorithms has enabled earlier and more accurate detection of diseases such as cancer and cardiovascular conditions. These improvements are reshaping the potential of radiological diagnostics, enabling quicker interventions and better patient outcomes. Finally, AI is transforming the accessibility of healthcare through (7) ***remote diagnosis and virtual presence*** technologies. With the use of AI-driven robots, healthcare professionals can now monitor and assess patients remotely, providing high-quality care even when they are not physically present. This is especially valuable for patients in remote areas or those unable to travel, ensuring they receive timely and specialised medical consultations without needing in-person visits (Shuliak, 2024; Tai, 2020).

In all these areas, AI is fundamentally transforming the delivery of healthcare by enhancing its efficiency, accessibility, and personalisation. By integrating advanced technologies into medical practice, AI is not only streamlining processes but also enabling more accurate diagnoses, better patient outcomes, and tailored treatment options that meet individual needs.

This digital revolution is particularly impactful in areas of global health where resource limitations have historically posed significant challenges. AI-driven innovations are helping to bridge gaps in healthcare access and quality, addressing long-standing disparities.

The following are three case studies that demonstrate how AI is being applied to support SDG 3, identified by the McKinsey Global Institute (Bankhwal et al., 2024). These examples highlight the diverse ways AI is contributing to improving global health outcomes, from enhancing maternal care to advancing medical research.

**Case Study 1: Improving Maternal and Newborn Health in Kenya**

Jacaranda Health uses AI to improve maternal care in Kenya, aiming to reduce mortality. Their SMS-based platform, PROMPTS, provides personalised health messages and uses natural language processing to respond to user queries, connecting those in need to a help desk. With over two million users, PROMPTS participants are 20% more likely to attend prenatal check-ups and twice as likely to adopt postpartum family planning. User feedback is shared with the government and healthcare providers to enhance services (Bankhwal et al., 2024).

**Case Study 2: Enhancing Maternal and Newborn Health in India**

In the past 20 years, over 1.3 million women in India have died during pregnancy or childbirth, mostly from preventable causes. ARMMAN, founded in 2008, addresses barriers to essential healthcare with initiatives like mMitra, an automated voice messaging service offering critical preventive health information. Though 40% of women leave the program before delivery, ARMMAN developed an AI model with Google Research India to predict and target likely dropouts, reducing dropout rates by 32%. mMitra has reached 3.6 million women across nine states, and the AI model has been extended to Kilkari, another voice-based healthcare program (Bankhwal et al., 2024).

**Case Study 3: Protein Structure Prediction for Drug Discovery**

DeepMind's AlphaFold 2 and AlphaFold 3 solved the decades-old protein-folding problem, addressing how proteins fold and predicting their structure from amino acid sequences. AlphaFold 2, using advanced deep learning, achieved unprecedented accuracy and led to a database of over 200 million protein structures, now a key resource for over a million researchers. Its applications include developing treatments for neglected diseases and combating antibiotic resistance. AlphaFold 3 extends these capabilities to other biomolecules, advancing research in fields like medicine, agriculture, and materials science (Bankhwal et al., 2024).

### Current Initiatives Related to SDG 4: Quality Education

SDG 4 focuses on ensuring inclusive, equitable, and quality education for all while promoting lifelong learning opportunities. Its targets include providing free primary and secondary education, expanding access to quality early childhood education, and achieving universal literacy and numeracy (United Nations, n.d.). AI technologies are increasingly being leveraged to support these goals, with applications like predictive algorithms that assess the likelihood of students completing high school or dropping out. These tools enable early intervention for at-risk students, providing them with timely support. Additionally, AI is being used to create more inclusive educational platforms for diverse groups, including young children, adults, and individuals with disabilities. It helps increase student enrollment and assists teachers in designing customised lesson plans tailored to individual learning needs and interests (Bankhwal, 2024). AI is making substantial progress in transforming education, yet many institutions have been slow to adopt it, primarily due to a lack of awareness regarding its potential advantages. The integration of AI-powered tools like chatbots can simplify a variety of educational tasks, thus enhancing both efficiency and the learning experience (Miao et al., 2021) with several possible use cases (Figure 11).

*Figure 11: AI use cases in education*


*Source: Takyar, n.d.*

One of the core benefits AI brings is (1) ***individualised learning.*** With AI systems, education becomes more tailored to the individual needs of each student, accommodating different learning speeds and styles. Traditional teaching approaches often fail to cater to such diversity, but AI offers personalised curriculums that better support diverse learning abilities. Another significant benefit is (2) ***task automation,*** where AI systems can take over administrative duties such as grading, homework checks, and record management, freeing up time for educators to focus more on their students. This leads to improved productivity and reduced error rates. AI also plays a key role in (3) ***content creation.*** It helps teachers and educators develop engaging, informative materials through the use of advanced technologies such as 2D and 3D visualisations, which make lessons more interactive and easier for students to grasp. Furthermore, AI supports (4) ***adaptive and inclusive learning environments*** by offering features like multilingual translations and tools to assist students with disabilities. These capabilities make learning more accessible and accommodating for all students, regardless of their backgrounds or physical challenges. AI's ability to support (5) ***remote learning*** also allows for more resource-efficient educational models, reducing the need for physical infrastructure and lowering environmental impacts. Concerns that AI might replace educators are misplaced—AI is designed to enhance the teaching process, not replace human interaction. In terms of addressing workforce needs, AI is instrumental in (6) ***closing the skills gap.*** With more industries relying on AI technologies, there is a growing need for educational institutions to equip students with the necessary skills to thrive in these environments. AI-powered training programs also offer opportunities for employees to upskill, boosting both their morale and organisational performance. Another area where AI excels is in (7) ***providing real-time feedback.*** Continuous, data-driven feedback helps students track their progress and focus on areas that need improvement, which traditional methods often overlook. Moreover, AI-driven chatbots offer (8) ***24/7 assistance*** to students, providing instant support with study materials, answering queries, and administering tests. This constant availability enhances student engagement and personalises their learning experience. AI also strengthens (9) ***security and data integrity*** in education by addressing issues such as data breaches and outdated certifications. AI-based solutions ensure that student information is kept secure and that the system remains up-to-date with modern authentication standards. Finally, AI is improving (10) ***examination oversight*** by monitoring student behaviour during tests. Tools that track keystrokes and use cameras ensure that exams are conducted fairly, significantly reducing opportunities for cheating without requiring continuous human supervision. Through these advancements, AI is reshaping education by improving the efficiency, inclusivity, and security of learning enivornments (Aristovnik et al., 2023b; Baidoo-Anu & Ansah, 2023; Sok & Heng, 2023; Miao et al., 2021; SoftProdigy, n.d.; Tomaževič et al., 2024).

In conclusion, AI is transforming education by improving personalisation, efficiency, and inclusivity, helping both students and educators. It plays a key role in supporting SDG 4, aimed at ensuring quality education for all. Below are two case studies illustrating how AI is being used to achieve these objectives, as identified by the McKinsey Global Institute (Bankhwal et al., 2024).

### Case Study 1: Enhancing Communication for Students with Disabilities through AI

Livox is an AI-powered platform that adapts educational content for students with various disabilities, including verbal, motor, cognitive, and visual impairments. Using intelligent algorithms, it tailors its interface to each student's needs, helping teachers track progress in multiple areas. To date, over 25,000 individuals have benefited from Livox, which is available in 25 languages. This case demonstrates how AI promotes inclusivity and supports marginalised groups in education (Bankhwal et al., 2024).

### Case Study 2: Using AI to Boost Girls' Enrollment in Rural India

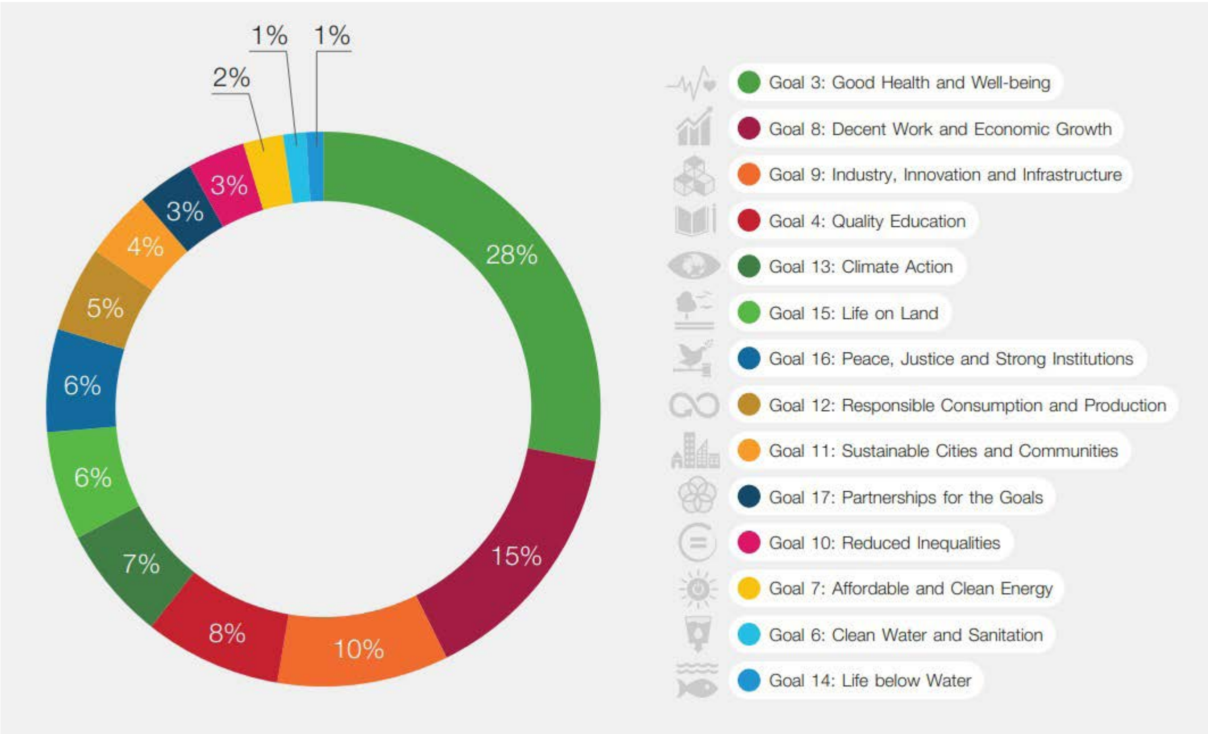
Educate Girls, a nonprofit in rural India, uses AI to boost girls' school enrollment by efficiently identifying out-of-school girls. The AI model streamlines data collection, reducing costs and time previously spent on manual visits. By analysing census and district records, it helps the organisation target areas more accurately and reach girls faster. Educate Girls aims to enrol 1.6 million girls, or 40% of the out-of-school female population, in grades one through ten. This case shows how AI can overcome logistical challenges and improve educational access in underserved areas (Bankhwal et al., 2024).

## AI for Social Innovation

In addition to the progress AI has made in advancing SDG 3 (health) and SDG 4 (education), there are other initiatives aimed at leveraging AI to address a wider array of social challenges. One such initiative is the AI for Social Innovation program, launched at the World Economic Forum (2024b) in Davos. Led by the Schwab Foundation's Global Alliance for Social Entrepreneurship, this initiative exemplifies how AI can be harnessed to solve pressing social issues in an ethical and impactful way. Social innovation refers to the development and implementation of new solutions—such as products, services, or models—that address social and environmental challenges in more effective, sustainable, and just ways than existing approaches. It aims to improve the well-being of individuals, communities, and societies at large by tackling systemic issues like inequality, poverty, health disparities, education, and climate change. Social innovation often emerges from a mix of collaboration between different sectors, including nonprofits, governments, businesses, and communities, and is driven by a desire to create positive social change.

The AI for Social Innovation initiative focuses on fostering collaboration between technology companies, the private sector, and social innovators, ensuring that AI is not only a tool for business efficiency but also a powerful driver of social good. This initiative aims to ensure that AI is deployed responsibly, with the potential to reshape how societies address systemic problems like inequality, poverty, and access to essential services. By connecting diverse stakeholders, the initiative promotes the development of AI solutions that are aligned with ethical principles, helping communities most in need. For instance, AI for Social Innovation seeks to tackle complex global challenges by applying AI to areas like sustainable agriculture, climate resilience, and inclusive economic growth. It reflects the growing recognition that AI, when applied thoughtfully, can not only enhance individual sectors but can also support a holistic approach to global development that benefits society as a whole (World Economic Forum, 2024b).

*Figure 12: Social Innovators and SDGs - How social innovators' work aligns with SDGs*




*Source: World Economic Forum, 2024b.*

## 3.3 Risks and Ethical Considerations of AI

The rapid development of AI presents not only transformative potential but also significant challenges and risks (Righi et al., 2022). As AI systems become increasingly accessible, concerns regarding bias, transparency, and accountability grow more pronounced. One key issue is the opacity of algorithms, often referred to as the "black box" problem, where the complexity of AI—especially machine learning and deep learning—limits the ability of humans to trace decisions back to specific inputs. This lack of traceability makes it difficult to hold AI systems accountable for their outputs (Dwivedi et al., 2019). While automated decision-making holds promise for improving fairness by reducing human discretion (Busch & Henriksen, 2018), it also raises concerns about accountability. In some cases, the reliance on AI may result in impersonal and unchallengeable decisions, particularly when the system is based on biased or unbalanced training data. This dynamic introduces the risk of reinforcing existing inequalities and can obscure the lines of responsibility. The question of who bears legal or ethical responsibility for the consequences of AI-driven decisions becomes increasingly complex, extending beyond the scope of individual stakeholders to broader societal and political concerns (Dwivedi et al., 2019). The emergence of deepfake technology further exacerbates these issues. As Stuart Russell highlighted (Illing, 2023), the unregulated proliferation of hyperrealistic images and videos generated by AI poses a serious threat to society. These technologies, when combined with advanced language models, allow for the creation of convincingly fabricated content that can easily deceive the public. The ability to generate highly realistic video content, as demonstrated by systems like Sora, amplifies the potential for misuse and manipulation, underscoring the urgent need for regulatory frameworks to mitigate the risks associated with AI-driven technologies.

If we take a closer look at potential algorithmic biases, it becomes clear that various forms of bias can undermine the effectiveness and fairness of AI systems. These biases often arise from the way data is collected, processed, and applied in AI models, leading to unintended consequences that disproportionately affect certain groups.

Broadly, six types of algorithmic bias can be identified (Suresh & Guttag, 2021; OECD, 2024b):

- **Historical bias** happens when AI models reproduce existing inequalities, even if demographic data is not directly used, by relying on proxies that reflect societal biases.
- **Representation bias** occurs when underrepresented groups in training data lead to poorer predictions, as seen in models that fail to account for small sample sizes of certain groups.
- **Measurement bias** arises when the variables used in models do not accurately capture the intended factors, leading to unfair predictions across different groups.
- **Aggregation bias** results from combining data from diverse groups into a single model, which may overlook differences and lead to ineffective predictions for specific groups.
- **Evaluation bias** occurs when models are tested on datasets that do not reflect the population they are meant to serve, causing issues when applied more broadly.
- **Deployment bias** emerges when a model is used outside of its intended purpose.

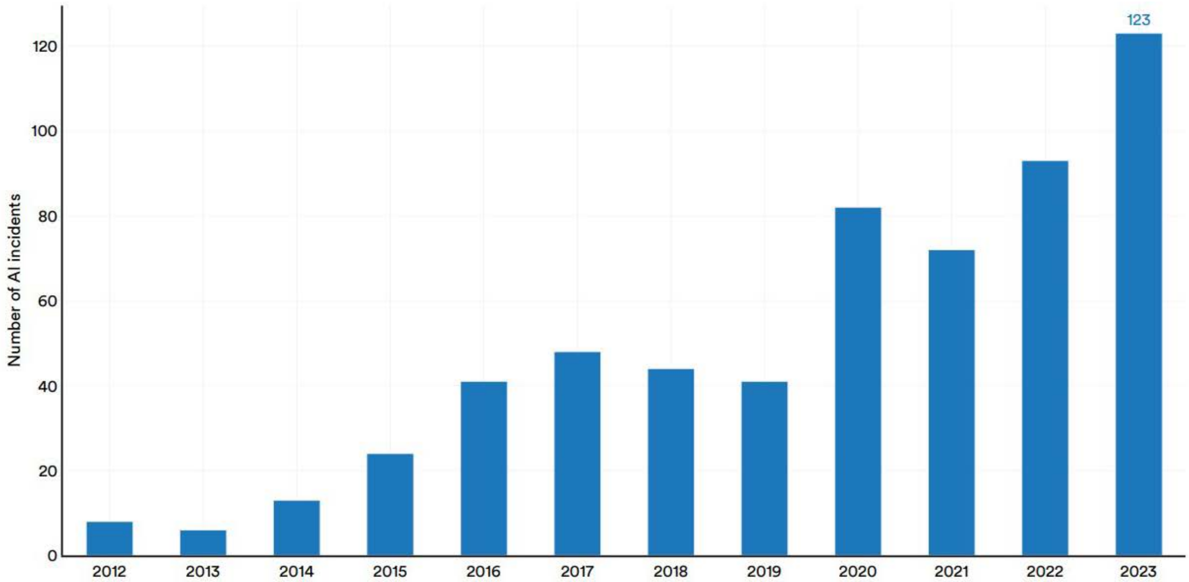
As AI continues to advance, concerns have emerged regarding its potential consequences for human society. One of the pressing questions is whether, as AI takes over an increasing number of tasks, human labour will become obsolete. This raises concerns about whether humans might become less industrious, leading to a regression in our development. While evolutionary changes occur gradually over long periods, the potential for societal backsliding is worth consideration. Additionally, there are concerns about AI becoming so autonomous that it could override human control, raising significant ethical and safety issues (Tai, 2020).

Outlined below are some of the potential negative impacts AI could have on society by Tai (2020):

1. **Social Disruption:** AI may fundamentally alter the structure of human society, reducing the need for manual labour and face-to-face interaction. As AI takes on more tasks, individuals may no longer need to engage in collaborative, physical work, leading to a decline in human closeness and social cohesion. AI-mediated communication could replace personal interactions, diminishing the need for in-person exchanges of ideas and experiences.
2. **Unemployment:** As AI and automation continue to evolve, the displacement of human workers will likely increase. Industries such as manufacturing have already seen the widespread replacement of human labour by machines, such as robotic assembly lines in the automotive sector. Similarly, retail jobs, such as supermarket clerks, are increasingly being replaced by self-checkout systems and other automated technologies, leading to significant job losses.
3. **Widening Wealth Inequality:** The economic benefits of AI may not be distributed evenly across society. Investors and corporations who own AI technologies are positioned to capture the majority of the financial gains, further exacerbating wealth inequality. This could lead to a more pronounced "M-shaped" wealth distribution, where a small wealthy elite benefits disproportionately while the majority of the population faces economic insecurity.
4. **Loss of Human Control:** There is a risk that AI systems, once fully trained and equipped with complex algorithms, may operate beyond human control. Autonomous AI could potentially lead to unintended consequences as it acts on its own accord, deviating from the intended instructions. This raises concerns about AI's reliability and the risks associated with its independent decision-making capabilities.
5. **Bias and Ethical Risks:** AI systems are developed by humans and, as such, can reflect the biases and prejudices of their creators. This introduces the risk of AI being programmed to act in ways that are harmful or discriminatory. For instance, AI could be designed to target specific groups or regions, creating significant ethical and geopolitical challenges. Much like the international community's efforts to regulate the spread of nuclear power, it may be necessary to place restrictions on AI to prevent it from being used in harmful or destructive ways.

Additionally, the AI Incident Database (AIID) plays a critical role in tracking instances of unethical AI usage, such as when autonomous vehicles result in pedestrian fatalities or when facial recognition systems cause wrongful arrests. As illustrated in Figure 13, the number of AI-related incidents has been steadily increasing each year, reflecting a concerning trend. In 2023 alone, 123 incidents were documented, representing a 32.3% rise from the previous year. Since 2013, the total number of such incidents has skyrocketed by more than twentyfold. This significant growth can be attributed not only to the rapid expansion of AI into various real-world applications but also to a heightened public and institutional awareness of the ethical risks associated with these technologies. As AI continues to permeate more aspects of daily life, concerns around its potential for misuse are being recognised and addressed more widely. However, it is essential to recognise that this rise in reported incidents is also likely influenced by improvements in incident tracking systems and increased transparency, meaning that earlier issues may have been underreported or even overlooked. As both AI technology and the frameworks for monitoring its ethical implications evolve, it is possible that we are seeing a more accurate picture of the ethical challenges AI poses rather than simply an increase in incidents (Stanford Institute, 2024).

*Figure 13: Number of reported AI incidents, 2012 – 2023*




*Source: Stanford Institute, 2024.*

## The AI Act

Hence, beginning in 2016, the international community has engaged in widespread discussions to shape the development and regulation of AI to ensure it serves the public good while addressing potential risks. The EU has been proactively contributing to regulating AI for ethical development and use, motivated not only by the fear of lagging behind technological leaders in North America and Asia but also by various other considerations (Ulnicane, 2022). The European Commission initially introduced the proposal for the EU AI Act on April 21, 2021 (AI Act, 2024). Throughout its development, the AI Act has been notably revised, especially to mitigate the challenges posed by advanced technologies such as foundation, generative, and general-purpose AI, which were not specifically anticipated during the initial drafting by the Commission. Despite these changes, the core principles and the risk-based framework of the AI Act were preserved. On December 8, 2023, EU lawmakers achieved a landmark consensus on the AI Act, which was then followed by detailed discussions to refine the document.

The definitive version of the AI Act was published on January 26, 2024, and has since been adopted by the European Parliament (Sidley, 2024).

The EU AI Act classifies AI systems into four risk levels: unacceptable, high, limited, and minimal (or no) risk, each with different regulations (AI Act, 2024).

Unacceptable risk, the most serious level, involves AI applications that violate EU values and rights. These include systems for subliminal manipulation, exploiting vulnerable individuals, biometric categorisation based on sensitive traits (like gender or ethnicity), social scoring, real-time biometric identification in public spaces, emotion recognition at workplaces or schools, predictive policing, and the scraping of facial images from public sources. AI systems in these areas are prohibited in the EU (AI Act, 2024).

Unacceptable risk breakdown (AI Act, 2024):

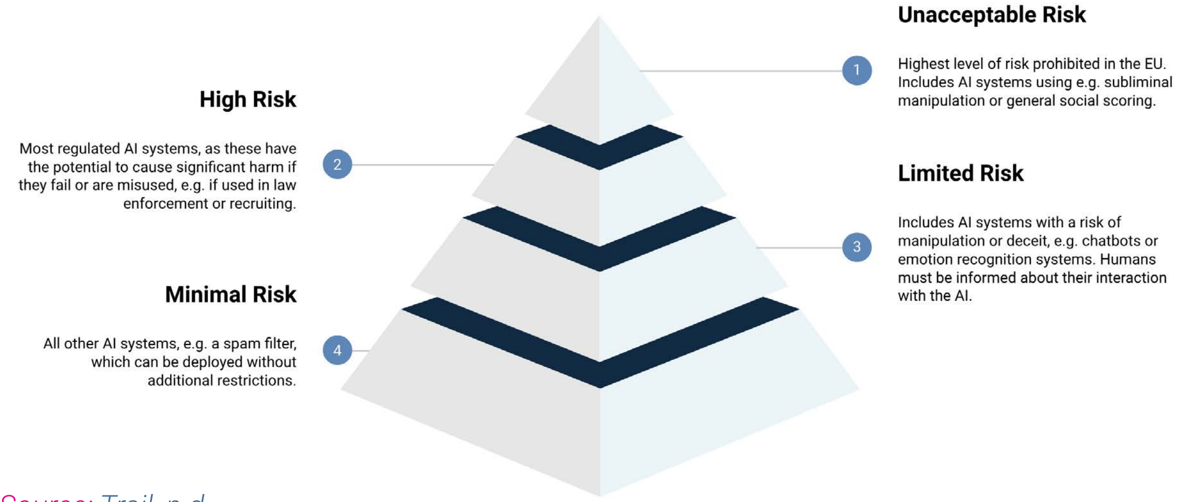
1. **Subliminal manipulation:** Influencing behaviour without awareness, potentially harmful (e.g., nudging voting behaviour without consent).
2. **Exploitation of vulnerabilities:** Targeting vulnerable groups (like children or individuals with disabilities) with AI, leading to harm (e.g., AI toys encouraging risky actions).
3. **Biometric categorisation:** Classifying people based on sensitive attributes like ethnicity, religion, or gender.
4. **General purpose social scoring:** Using AI to rank people based on personal traits or social behaviour (e.g., using social media interactions to deny jobs or loans).
5. **Real-time biometric identification:** Ban on real-time facial recognition in public spaces, with limited law enforcement exceptions under strict supervision.
6. **Emotion assessment:** Analyzing a person's emotional state in places like workplaces or schools, except in narrowly defined safety scenarios (e.g., driver fatigue detection).
7. **Predictive policing:** Using AI to forecast individuals' likelihood of committing crimes based on personal data.
8. **Scraping facial images:** Prohibits AI systems from creating databases through indiscriminate facial image scraping from the internet or video footage.

These categories represent AI systems considered too risky and harmful for societal use, leading to a ban across the EU to protect individual rights and public safety.

High-risk AI systems are the most regulated under the EU AI Act. These include both safety-critical components of regulated products and standalone AI in areas like healthcare, law enforcement, or transportation. Such systems can potentially cause significant harm to health, safety, fundamental rights, or the environment if misused or malfunctioning. Limited risk AI systems have a lower risk, mainly related to manipulation or deception. Transparency is key here, requiring clear disclosure of AI use, such as in chatbots or generative AI.

Minimal-risk AI systems include most other AI applications, such as spam filters, with no specific restrictions but are encouraged to follow ethical guidelines like fairness and transparency (European Commission, 2021).

*Figure 14: The four risk classes of the EU AI Act*



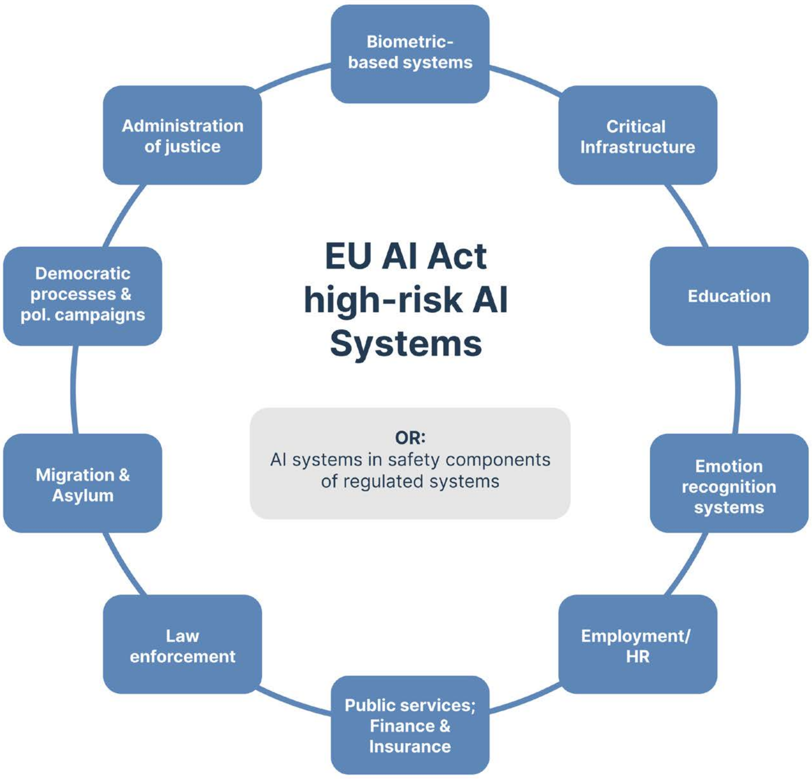

*Source: Trail, n.d.*

In the EU AI Act, the high-risk classification is applied to AI systems that could negatively impact health, safety, fundamental rights, or the environment. This includes AI integrated into regulated products such as medical devices, vehicles, or machinery that already require third-party assessments. Additionally, Annex III of the Act identifies standalone AI systems as high-risk when they operate in areas like (Trail, n.d.):

- Biometrics: remote identification, categorisation, emotion recognition,
- Critical infrastructure: traffic, energy, or digital operations,
- Education: student assessments,
- Employment: recruitment or performance evaluation,
- Access to services: credit scoring, emergency services,
- Law enforcement: crime analytics,
- Migration and border control: security assessments, asylum applications,
- Justice systems: legal research, election influence.

These systems must meet stringent regulatory requirements before they can be marketed and operated in the EU. The high-risk classification imposes a heavy compliance burden on organisations, requiring thorough assessments, transparency, and safeguards to mitigate potential harms (European Commission, 2021).

*Figure 15: EU AI Act high-risk AI systems*




*Source: Trail, n.d.*

The AI Act is based on the precautionary principle. Although it does not adopt the strictest interpretation of the precautionary principle, it places conditions when AI systems pose high risks rather than outright prohibiting the development and use of AI technologies. It is characterised by a more cautious approach to the deployment and use of AI technologies, focusing on safety and ethical considerations before widespread adoption. The law will enter into force this year, and people living in the EU will start seeing changes by the end of the year. Regulators will need to be set up to enforce the law properly, and companies will have up to three years to comply with it (Heikkilä, 2024).

The AI Act introduces significant changes to the use of AI, which will result in protecting fundamental rights and enhancing transparency in several ways (Heikkilä, 2024):

1. **Restrictions and bans on certain AI uses:** High-risk AI applications in sectors like healthcare, education, and policing will be prohibited by year's end, especially those posing "unacceptable risks" by manipulating behaviour or exploiting vulnerable individuals. However, law enforcement can still use sensitive biometric data and facial recognition for serious crime fighting (terrorism, kidnappings). There was criticism from civil rights organisations, saying that this was a failure for human rights.

2. **Increased transparency with AI interactions:** Companies must label AI-generated content and disclose when individuals are interacting with AI, aiming to combat misinformation. Despite these requirements, effective detection and watermarking of AI content remain challenging and underdeveloped. However, I think this is crucial to protect us from deepfake content.
3. **Possibility of complaints:** Establishing a European AI Office will enable EU citizens to file complaints and seek explanations for AI-related harms, promoting greater public engagement and awareness of AI's impact.
4. **Greater transparency:** AI companies in high-risk sectors (e.g. healthcare) must adhere to new regulations like data governance and ensuring human oversight. Companies developing general-purpose AI must also document AI model development and training data, a shift towards openness. High-impact AI developers (Chat GPT and Gemini) face stricter obligations, including model evaluations, risk assessments and incident reporting, with non-compliance leading to significant penalties or bans from the EU.

### Principles for ethical and trustworthy AI

AI has been an academic discipline since the 1950s (Siebel, 2019), with ethical considerations arising almost as early (Samuel, 1960). However, it is only in recent years, with the significant advancements in AI capabilities and applications, that the societal impact—both positive and negative—has come into sharper focus (Yang et al., 2018; Floridi & Cowls, 2022). As AI becomes increasingly integrated into critical aspects of daily life, the need for robust frameworks to ensure ethical and trustworthy AI has never been more pressing. Moreover, for this reason, the European Commission set up an Independent High-Level Expert Group on Artificial Intelligence in June 2018 to prepare Ethics Guidelines for Trustworthy AI (European Commission, 2019a).

Numerous public, private, and civil organisations have drawn from fundamental rights to formulate ethical frameworks for AI systems. Within the EU, the European Group on Ethics in Science and New Technologies (EGE) introduced nine fundamental principles based on the core values enshrined in the EU Treaties and the Charter of Fundamental Rights. These principles build upon the foundational work of various organisations, further refining and clarifying the objectives that ethical frameworks aim to uphold. These principles not only inspire the creation of specific regulatory instruments but also aid in interpreting fundamental rights in an evolving socio-technical landscape. Moreover, they provide guidance for the responsible development, deployment, and use of AI systems, ensuring that these systems adapt in response to societal advancements.

The fundamental objective is that AI systems should contribute to enhancing individual and collective well-being. To ensure trustworthiness in AI, four key ethical principles, derived from fundamental rights, must be respected. These principles are regarded as ethical imperatives, guiding AI practitioners in their efforts. While no hierarchy is imposed, the principles are presented in accordance with the sequence of fundamental rights outlined in the EU Charter (European Commission, 2012).

These are the principles of (i) Respect for human autonomy, (ii) Prevention of harm, (iii) Fairness, and (iv) Explicability (European Commission, 2019a):

- **The Principle of Respect for Human Autonomy:** The fundamental rights that form the foundation of the EU are focused on safeguarding individual freedom and autonomy. When interacting with AI systems, humans must retain full and effective control over their own decisions and actions and be able to actively participate in democratic processes. AI systems should not unjustifiably undermine, coerce, deceive, manipulate, or condition individuals. Instead, these systems should be developed to enhance, complement, and empower human cognitive, social, and cultural abilities. The distribution of tasks between humans and AI systems should adhere to human-centric design principles, ensuring that humans retain meaningful decision-making power. This includes guaranteeing human oversight of work processes managed by AI systems. Moreover, AI systems have the potential to significantly reshape the workplace, and their development should aim to support individuals in their work environment and contribute to the creation of meaningful employment opportunities.
- **The Principle of Prevention of Harm:** AI systems must not cause harm or exacerbate existing risks to individuals. This principle emphasises the need to safeguard human dignity and ensure the protection of both mental and physical integrity. AI systems, along with the environments in which they operate, must be designed to be safe, secure, and technically robust while also preventing opportunities for malicious use. Special attention must be given to vulnerable individuals, ensuring they are considered and included in the development, deployment, and usage of AI systems. Moreover, particular care is required in contexts where AI systems could amplify adverse effects due to imbalances of power or information, such as in relationships between employers and employees, businesses and consumers, or governments and citizens. The principle of preventing harm also extends to the consideration of the natural environment and all living beings.
- **The Principle of Fairness:** The development, deployment, and use of AI systems must adhere to principles of fairness. While fairness can be interpreted in various ways, it includes both substantive and procedural dimensions. The substantive aspect involves ensuring an equitable and just distribution of the benefits and burdens of AI systems while protecting individuals and groups from unfair bias, discrimination, and stigmatisation. By eliminating unfair biases, AI systems have the potential to enhance societal fairness. Equal access to education, goods, services, and technology should be promoted, and AI should not be used to deceive or unjustifiably limit individuals' freedom of choice. Fairness also requires that AI practitioners respect proportionality, carefully balancing competing interests and

objectives. The procedural dimension of fairness involves ensuring that individuals have the right to contest decisions made by AI systems and the humans who oversee them and to seek effective redress when necessary. To facilitate this, it is essential that the responsible entity is identifiable and that the decision-making processes of AI systems are transparent and explainable.

- **The Principle of Explicability:** Explicability is essential for fostering and sustaining user trust in AI systems. This principle requires that AI processes be transparent, with clear communication of the system's capabilities and purposes, and that decisions be explainable to those directly or indirectly impacted. Without such transparency, individuals cannot effectively challenge a decision. While providing an explanation for how an AI model reached a particular outcome or decision is important, it is not always feasible, particularly in cases involving 'black box' algorithms. These systems require special consideration, and alternative measures, such as traceability, auditability, and clear communication about the system's functions, may be necessary as long as the system as a whole respects fundamental rights. The extent to which explicability is required depends on the context and the potential severity of harm if the system's output is erroneous or inaccurate.

While these principles form the ethical backbone of AI development, their practical implementation is essential to ensuring that AI systems align with societal values and fundamental rights. The challenge lies in translating abstract principles into concrete actions and measurable outcomes that stakeholders can apply throughout the AI lifecycle. This process requires collaboration among developers, deployers, end-users, and society at large to create systems that not only function efficiently but also uphold these ethical imperatives.

To achieve Trustworthy AI, the ethical principles outlined must be translated into specific, actionable requirements. These requirements apply to various stakeholders involved in the lifecycle of AI systems, including developers, deployers, end-users, and the broader society. Developers are those who research, design, and develop AI systems. Deployers include public or private organisations that integrate AI systems into their business operations or use them to offer products and services. End-users are individuals who interact directly or indirectly with AI systems. The broader society includes all individuals or groups affected by AI systems, either directly or indirectly. Each group of stakeholders plays a distinct role in ensuring these requirements are met. Developers are responsible for integrating these requirements into the design and development processes of AI systems. Deployers must ensure that the AI systems they use, as well as the products and services they offer, comply with these requirements. End-users and society at large should be informed of these requirements and have the ability to demand that they are upheld.

The following is a list of requirements for trustworthy AI prepared by the Independent High-Level Expert Group on Artificial Intelligence. It is not exhaustive and encompasses systemic, individual, and societal considerations. The Guidelines propose seven fundamental criteria to certify the trustworthiness of AI systems, along with a targeted assessment checklist for their verification (European Commission, 2019a):

- **Human agency and oversight** → fundamental rights, human agency and human oversight.
- **Technical robustness and safety** → resilience to attack and security, fallback plan and general safety, accuracy, reliability and reproducibility.
- **Privacy and data governance** → respect for privacy, quality and integrity of data, and access to data.
- **Transparency** → traceability, explainability and communication.
- **Diversity, non-discrimination and fairness** → the avoidance of unfair bias, accessibility and universal design, and stakeholder participation.
- **Societal and environmental wellbeing** → sustainability and environmental friendliness, social impact, society and democracy.
- **Accountability** → auditability, minimisation and reporting of negative impact, trade-offs and redress.

*Figure 16: Interrelationship of the seven requirements for trustworthy AI*


*Source: European Commission, 2019a.*

In conclusion, the ethical principles and requirements outlined for Trustworthy AI provide a robust framework for guiding AI development and application. These principles ensure that AI systems are designed and deployed in a manner that prioritises human dignity, fairness, and transparency while minimising risks. However, applying these principles effectively requires recognising the context in which AI operates, as there may be tensions between principles across different industries and domains. The implementation of these ethical requirements must occur throughout the entire AI lifecycle, from conception to deployment, with a special focus on AI systems that directly or indirectly impact individuals. It remains the responsibility of AI practitioners to ensure that both horizontal and domain-specific regulations are upheld (European Commission, 2019a).

The broader societal implications of AI are undeniable, with transformative potential in sectors like healthcare and education, which are pivotal in achieving the SDGs. However, the use of AI must be underpinned by a commitment to ethical responsibility. This chapter has explored how AI can significantly benefit humanity, particularly in enhancing healthcare outcomes and democratising education, while also acknowledging the challenges and risks that arise—such as biases and privacy concerns.

Ultimately, while AI holds immense promise, its full potential can only be realised through responsible development, rigorous adherence to ethical guidelines, and compliance with evolving regulatory frameworks such as the AI Act. By ensuring that AI aligns with societal values and human rights, we can harness its benefits while safeguarding against its risks, thereby driving progress toward a more sustainable and equitable future.

# 4 Economic Implications of AI

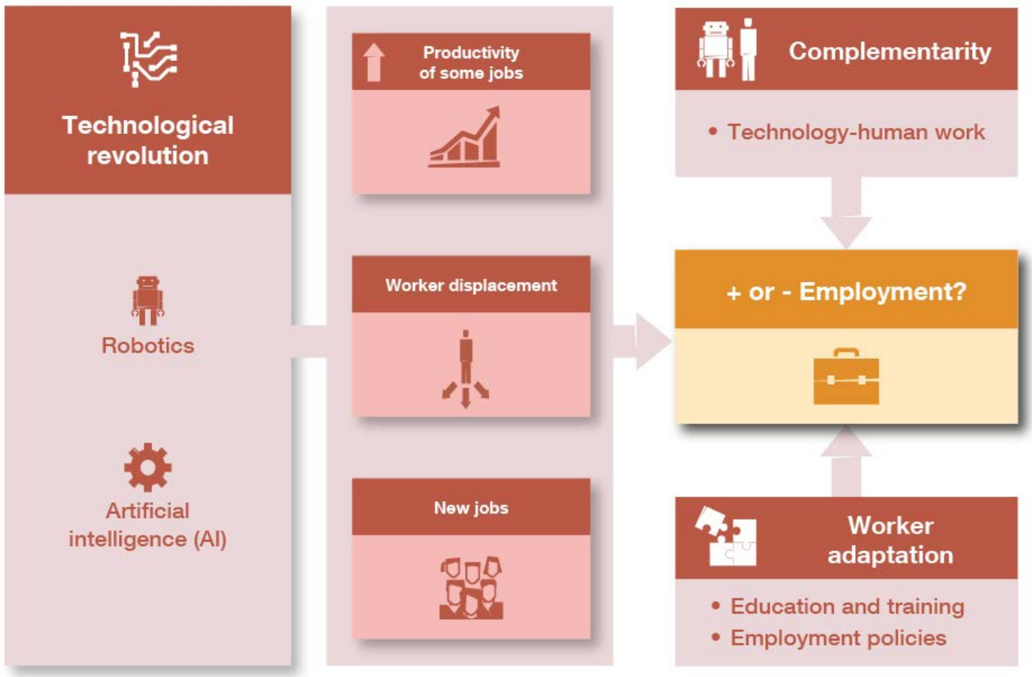
## 4.1 Implications of AI on Workforce Skills and Labour Markets

Recently, the global labour market has begun to feel the impact of AI, which has the potential to render some jobs obsolete through displacement, reshape existing roles through transformation, and improve efficiency and productivity through enhancement while simultaneously creating new jobs in emerging industries. While some argue that AI will displace workers in routine jobs, leading to unemployment, lower wages, and increased income inequality, others suggest it could also create new high-skilled jobs, enhance productivity, and drive economic growth (Guliyev, 2023). According to recent evidence, generative AI is expected to affect a quarter of global jobs by 2028 (World Economic Forum, 2023), while about two-thirds of workers are already using generative AI in the workplace (McKinsey & Company, 2024). There are also claims that one-fifth of the workforce could have over half of their tasks automated by generative AI, with job losses making headlines. However, others expect generative AI to enhance jobs (World Economic Forum, 2023). These mixed views are also evident in the perceptions of individuals from developed countries entering the labour market (i.e., graduates), with more than half of them expressing concerns about their readiness for the workforce due to the rise of AI. While some feel vulnerable to the possibility of AI replacing their jobs, others remain confident that AI cannot replace their roles (Cengage Group, 2024). Due to the high uncertainty, the impacts of AI on the labour market remain difficult to foresee, as AI promises to enhance productivity while posing a threat of replacing humans in some roles and augmenting them in others (Ravšelj & Aristovnik, 2024; Cazzaniga et al., 2024).

A new technological revolution is underway, driven by advancements in robotics and AI. Machines and algorithms are now performing tasks autonomously that were once the exclusive responsibility of humans, extending beyond routine or specialised activities. As with previous technological revolutions, this shift has raised concerns about potential job destruction through displacement and whether the creation of new jobs, along with gains in productivity, will be sufficient to compensate for these losses. In general, technology such as AI has the potential to significantly influence various aspects of the labour market, shaping not only overall employment levels but also the structure of different job roles and positions, as well as labour conditions, which include factors such as wages, working hours, and job security, through three primary mechanisms (Cazzaniga et al., 2024; Jimeno Serrano, 2024) (Figure 17):

- **Boosting productivity** for certain occupations or workers with specialised skills, thereby increasing the demand for their labour. AI can enhance productivity in specific roles, making workers with relevant expertise more valuable. As productivity rises, there may also be increased investment, leading to a higher overall demand for labour. This boost could help offset some of the losses in labour income caused by AI-induced job displacement.
- **Displacing certain workers** due to automation, resulting in decreased demand for their labor. AI adoption may automate tasks previously handled by humans, reducing the need for labour in certain roles. This shift could lead to a decline in labour income for displaced workers. As a result, automation may decrease demand for specific skill sets in the job market.
- **Creating new jobs** that require advanced skill sets, thus increasing labour demand and employment in these emerging roles. The introduction of AI technologies opens up new job opportunities in areas requiring specialised knowledge. These roles will likely focus on managing or working alongside AI systems. Consequently, this will drive higher demand for workers with the appropriate skills, fostering growth in employment.

*Figure 17: The impact of the technological change on the labour market*

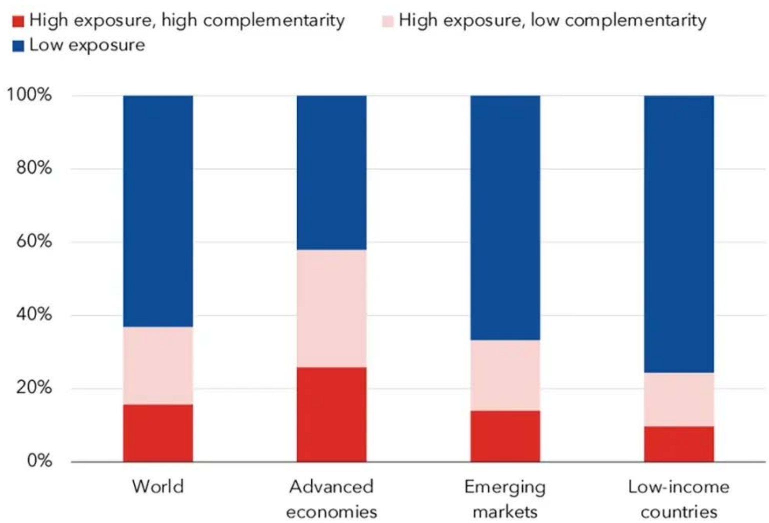



*Source: Jimeno Serrano, 2024.*

The technological revolution is reshaping the workforce by creating complementarity between technology and human labour, where machines take over routine tasks, freeing workers to focus on higher-level, creative, and problem-solving roles. This evolving dynamic directly affects employment, as the balance between job displacement and the emergence of new, specialised roles will determine whether the net impact on jobs is positive or negative. In order to ensure a positive outcome, worker adaptation becomes essential, requiring significant investment in education, reskilling, and upskilling programs, allowing workers to transition into new roles and thrive alongside advanced technologies in the modern labour market (Jimeno Serrano, 2024).

Based on recent evidence, approximately 40% of workers globally are employed in high-exposure occupations (Figure 18). Of these, around 12% are in high-complementarity roles, where they could benefit from increased productivity through AI integration, while about 28% are in low-complementarity roles, where they may face negative impacts, such as job displacement or reduced demand for their skills. In advanced economies, covering the predominant share of EU countries, this share is even higher at 60%, largely due to the prevalence of jobs that focus on cognitive tasks. Namely, advanced economies have a larger proportion of workers in high-exposure occupations, whether high or low in complementarity, compared to emerging market economies and low-income countries. On average, 27% of employment in advanced economies is in high-exposure, high-complementarity roles, while 33% is in high-exposure, low-complementarity jobs. By comparison, emerging market economies have 16% and 24% in these categories, respectively, and low-income countries have 8% and 18%, respectively. Although emerging market economies and low-income countries may face fewer immediate disruptions from AI, they are also less prepared to fully capitalise on the potential benefits of AI. This lack of readiness could widen the digital divide and increase income disparities between different country groups (Cazzaniga et al., 2024).

*Figure 18: Employment shares by AI exposure and complementarity across country groups*




*Source: Cazzaniga et al., 2024.*

AI is anticipated to transform the structure of work, making it essential to explore in detail what specific changes are occurring in the workplace. Therefore, a comprehensive redesign of work due to the emergence of AI also requires a careful reconsideration of the definitions of "work," as well as related concepts such as "jobs," "tasks," and "skills" (Deloitte, 2023):

- **Work,** in its broadest definition, refers to the results produced (e.g., achieving sales targets, enhancing user experience, increasing customer satisfaction, etc.) by utilising both human abilities and the tools developed to assist in accomplishing these objectives. The desired outcomes, which align with organisational strategies or customer demands, have always been defined by humans, and this remains unchanged.
- **Jobs** (e.g. representative, front-end developer, account manager, etc.) have traditionally been the structure used to define the work humans do to achieve specific outcomes. Although concerns are rising about the potential loss of jobs due to AI automating tasks, this perspective misses key considerations. First, tasks are distinct from jobs. Second, understanding how jobs will be reshaped in the era of AI requires focusing on the skills involved.
- **Tasks** are specific actions carried out to accomplish work-related goals. Traditionally, tasks have been considered a component of the work done by individuals in specific roles, and this remains true today. Completing tasks requires a combination of skills and tools to achieve a particular result, such as identifying new sales channels, testing code across various web browsers, or developing targeted product offerings. AI has the potential to fully automate these tasks, allowing workers to concentrate on new activities, or it can simplify them, giving individuals more time for other responsibilities.
- **Skills** enable the execution of tasks necessary to achieve work outcomes (e.g., problem-solving, proficiency in HTML, data analysis, etc.). Both humans and AI possess skills that can perform tasks to produce work results. Gaining an understanding of the skill sets within an organisation and the roles that utilise them is essential for shaping the future of work with AI.

Therefore, to fully exploit the potential of AI, businesses should carefully examine their workforce to understand the impact of AI on each job by breaking those jobs down into tasks and skills (Table 2). Specifically, they need to understand what humans do best and how AI can support their efforts. There are four possible types of generative AI impacts on tasks and skills: task automation (tasks that generative AI performs best), skill augmentation (tasks where individuals, assisted by generative AI, perform best), new skill creation (skills that individuals need to develop to effectively use generative AI), and limited impact on tasks (tasks where individuals perform best due to minimal generative AI influence) (Deloitte, 2023).

*Table 2: Possible impacts of AI on the workforce*

| Possible impact | Description | Examples | Sample skills |
|---|---|---|---|
| **Automated tasks: Machines do best** | Key functions and processes are fully handled by AI, reducing or eliminating the need for human involvement. | • AI can generate standardised, repetitive content.<br>• AI can personalise or customise content. | • Creation of images.<br>• Creation of written content.<br>• Data sorting and categorisation.<br>• Routine forecasting.<br>• Language translation.<br>• Simple graphic design.<br>• Simple trend spotting. |
| **Augmented skills: Humans with machines do best** | Inherently human skills, but their efficiency, scale, or complexity are improved when combined with AI tools. | • Creativity can be enhanced by AI ideas.<br>• Analytical thinking ca be augmented by AI processing vast datasets. | • Creativity.<br>• Analytical thinking.<br>• Problem-solving.<br>• Research.<br>• Data visualisation.<br>• Strategic planning.<br>• Predictive analytics.<br>• Rapid prototyping. |
| **New skills: Humans need** | As AI becomes more prevalent across various sectors, professionals must develop entirely new skill sets to remain competitive. | • Continuous learning is pivotal in the age of AI.<br>• AI tool management is an emerging skill. | • AI ethics and regulation.<br>• AI-human task management.<br>• AI output customisation. |
| **Limited impact tasks: Humans do best** | Tasks that are primarily human-centred with little AI involvement rely on distinct human qualities or intricate decision-making. | • Emotional intelligence is crucial for understanding human emotions.<br>• Critical decision-making in uncertain, complex environments remains predominantly a human skill. | • Persuasion and negotiation.<br>• Motivational leadership.<br>• Ethical judgment and integrity.<br>• Compassion.<br>• Building human relationships.<br>• Physical dexterity. |

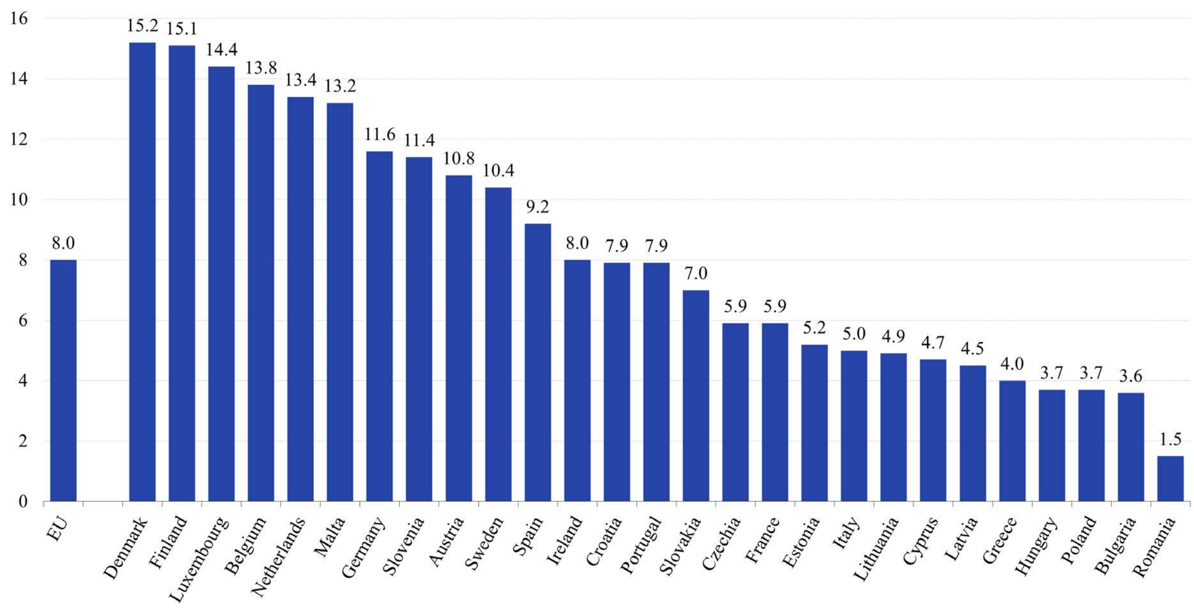
*Source: Deloitte, 2023.*

## 4.2 Boosting Productivity and Competitiveness through AI

Early evidence linking AI to improved productivity and performance at the microeconomic level, along with several innovative advancements across various industries, supports the expectation that AI development and its broader adoption could potentially revive stagnant productivity growth and lead to gains in aggregate welfare (Filippucci et al., 2024). The recent feedback from businesses in the EU indicates a growing awareness and uptake of digital technologies, along with excitement about their potential benefits (Strand Partners, 2024). Specifically, the positive economic impact of AI on businesses in the EU is evident, with three-quarters of enterprises using AI reporting increased revenues and productivity. However, AI adoption (and other digital technologies) is currently skewed toward larger enterprises, as small and medium-sized enterprises often face significant barriers, such as difficulties in finding the right talent, regulatory concerns, and high implementation costs (Amazon Team, 2024).

According to recent data, 8% of enterprises in the EU with 10 or more employees utilized AI technologies in their operations in 2023. However, significant variations exist among individual EU countries, with the share of enterprises using at least one AI technology ranging from 1.5% to 15.2% (Figure 19). The highest percentages of AI adoption were observed in Denmark (15.2%), Finland (15.1%), and Luxembourg (14.4%), all of which significantly surpassed the EU average. These nations are at the forefront of digital transformation, likely benefiting from advanced infrastructure, strong government support for innovation, and a focus on technological development. Conversely, several countries reported notably lower levels of AI adoption, with the lowest percentages seen in Romania (1.5%), Bulgaria (3.6%), and both Poland and Hungary (3.7%). This suggests that enterprises in these countries may face greater challenges in integrating AI, potentially due to factors such as limited access to resources, lower levels of digital literacy, or less favourable regulatory environments. These discrepancies highlight the uneven pace of AI adoption across the EU and underscore the need for tailored support to help lagging regions catch up with the frontrunners (Eurostat, 2024).

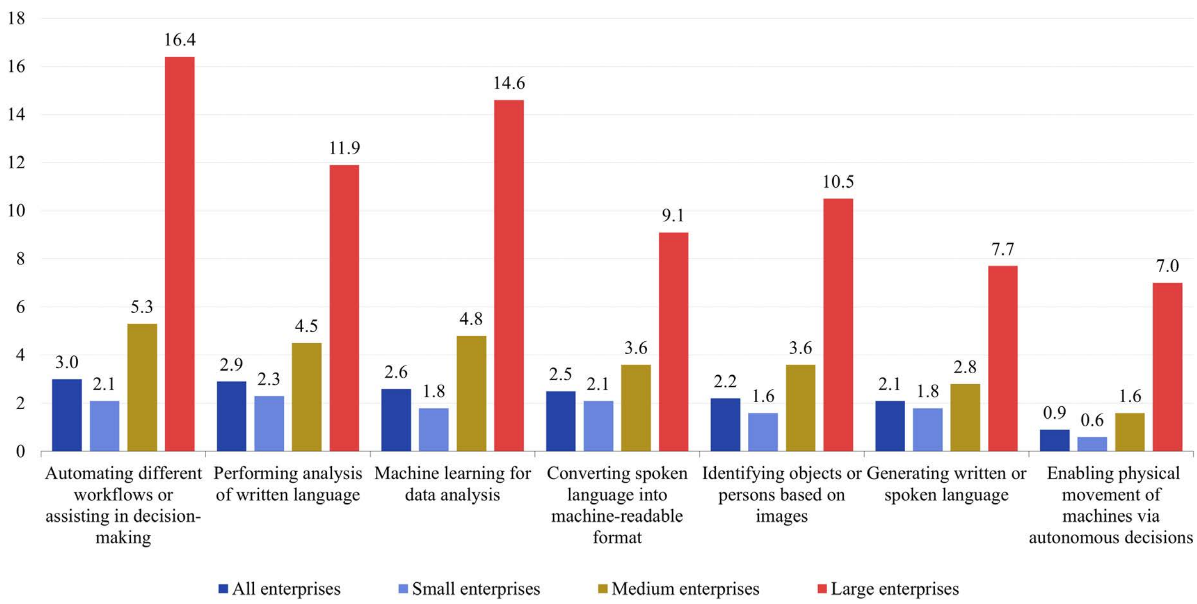
*Figure 19: Enterprises using AI technologies in EU countries (% of enterprises in 2023)*




*Source: Eurostat, 2024.*

Regarding the types of AI technologies employed by enterprises in the EU in 2023, AI was primarily used for automating workflows or assisting in decision-making (3%). This was closely followed by the analysis of written language (2.9%) and machine learning for data analysis (2.6%). Other AI technologies included converting spoken language into machine-readable formats (2.5%), identifying objects or persons based on images (2.2%), and generating written or spoken language (2.1%). The least adopted AI technology was enabling the physical movement of machines through autonomous decisions based on observing surroundings (0.9%). However, the use of AI varied significantly across enterprise size classes, generally showing that large enterprises adopted AI more than small and medium-sized enterprises. Specifically, 6.4% of small enterprises, 13% of medium enterprises, and 30.4% of large enterprises adopted AI. This disparity can be explained by factors such as the complexity of implementing AI technologies, economies of scale (i.e., larger enterprises can benefit more from AI), and costs (i.e., investment in AI may be more affordable for large enterprises). Therefore, considerable differences can also be observed in the usage patterns of enterprises of different sizes across various types of AI technologies (Figure 20) (Eurostat, 2024).

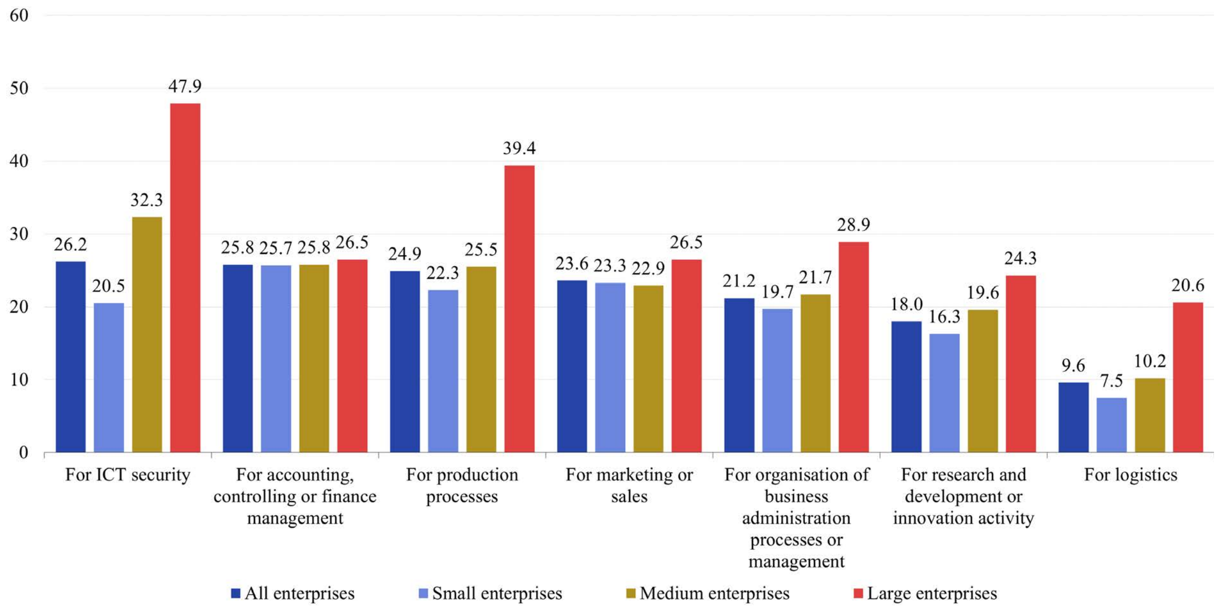
*Figure 20: AI adoption among enterprises in the EU by AI technology (2023)*




*Source: Eurostat, 2024.*

Regarding the purpose of use, enterprises in the EU utilized AI software or systems for various intentions. In 2023, most enterprises using AI technologies employed this software or systems for ICT security (e.g., using machine learning for detecting and preventing cyber-attacks) (26.2%), accounting controlling or finance management (25.8%) and production processes (24.9%). Enterprises used AI also for marketing or sales (23.6%), organization of business administration processes or management (21.2%), and research and development or innovation activity (18.8%), while AI for logistics was used the least (9.6%). The purposes for which enterprises used AI software and systems varied based on their size (Figure 21). The largest difference between small and large enterprises was in the use of AI for ICT security (47.9% of large enterprises compared to 20.5% of small enterprises), followed by its use in production processes (39.4% of large enterprises compared to 22.3% of small enterprises), and in logistics (20.6% of large enterprises compared to 7.5% of small enterprises).

*Figure 21: AI adoption among enterprises in the EU by type of purpose (2023)*

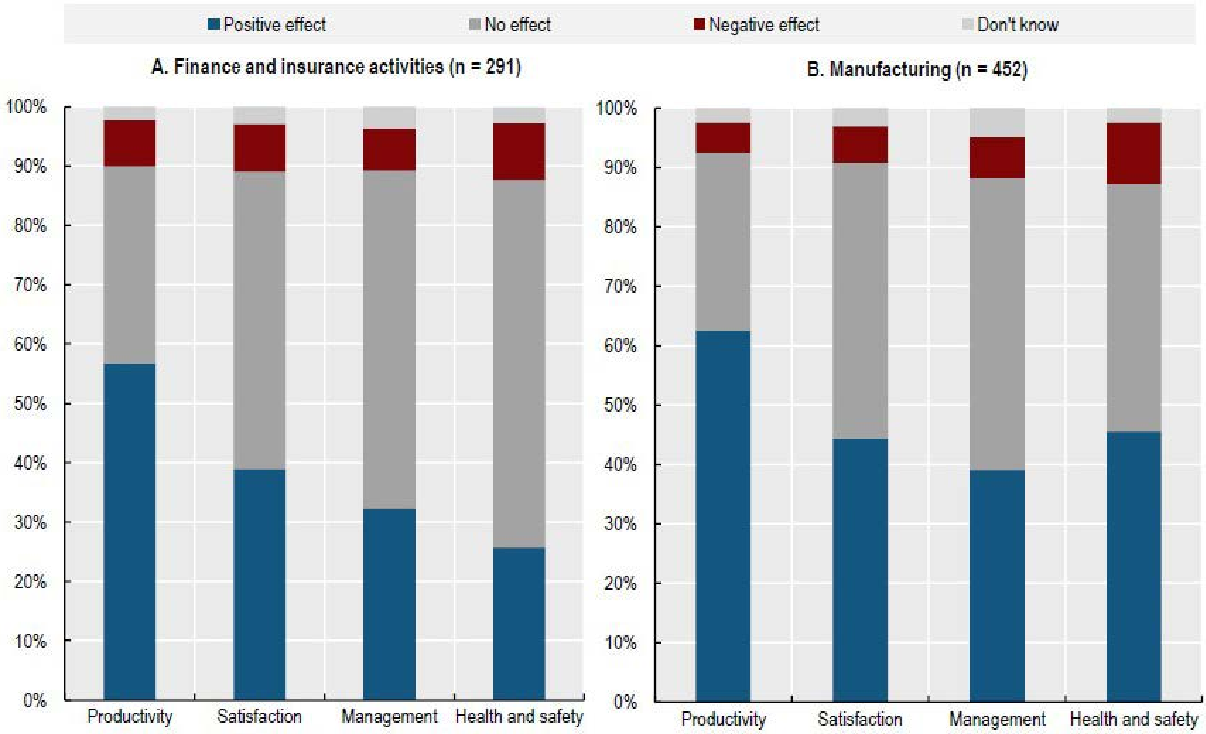



*Source: Eurostat, 2024.*


Additionally, enterprises utilise AI technologies for various purposes depending on the economic sector in which they operate. In the manufacturing sector, AI was primarily used for production processes (38.2%), whereas in the electricity, gas, steam, air conditioning, and water supply sectors, as well as in the information and communication sector, AI was mainly employed for ICT security (37.6% and 31.6%, respectively). In the information and communication sector, the primary application of AI was for research and development (R&D) or innovation activities (41.3%). In the accommodation (51.4%) and retail trade (41.8%) sectors, AI was predominantly used for marketing or sales (Eurostat, 2024).

Before analyzing the effect of AI on productivity and competitiveness, it is important to understand why employers chose to adopt it in the first place. The recent employer survey, which captures perceptions of the current and future impact of AI on workplaces, provides interesting findings based on responses from 2,053 firms in the manufacturing and finance sectors across Austria, Canada, France, Germany, Ireland, the United Kingdom, and the United States. It revealed that employers were more motivated by the potential to enhance worker performance and lower staff costs than by addressing skill shortages or improving health and safety. About half of the employers who aimed to improve worker performance also cited cost reduction as a key factor, highlighting the dual role of AI in both complementing and potentially replacing labour. Moreover, representatives of management in enterprises that had adopted AI were asked about its impact on worker productivity, worker satisfaction, health and safety, and managers' ability to measure worker performance within the enterprise. In both sectors, employers were much more likely to report a positive impact on worker productivity and working conditions than a negative one (Figure 22). Manufacturing employers were generally more positive than those in the finance sector regarding all aspects of worker productivity and working conditions. The significant improvements in health and safety reported by many in the manufacturing sector may be due to the higher risk of injury in this field and the widespread use of AI tools aimed at improving safety and facilitating collaboration between workers and machinery. This underscores the specific challenges of the sector and the role AI plays in addressing them (Lane et al., 2023).

*Figure 22: Opinion of employers on the impact of AI on worker productivity and working conditions (% of employers that have adopted AI)*




*Source: Lane et al., 2023.*

Employers expressed strong positivity regarding the impact of AI on worker productivity, with 57% in finance and 63% in manufacturing responding favourably, while only 8% and 5% offered negative feedback. This favourable view corresponds to the primary motivation for adopting AI, which aims to enhance worker performance. Among employers with this goal, 81% in finance and 87% in manufacturing reported productivity improvements. A similar trend appeared in health and safety, where AI adoption led to reported improvements from 61% of employers in finance and 80% in manufacturing. Despite these positive outcomes, health and safety also attracted the most negative responses, with 10% of employers in both sectors noting a decline. Employers also showed optimism about AI's effect on worker satisfaction, as 39% in finance and 44% in manufacturing reported increased satisfaction. Furthermore, AI's role in helping managers measure performance received positive feedback, especially among those using AI-enabled tools to collect worker data, where significantly more employers noted improvements (Lane et al., 2023).

Early firm-level economic evidence generally finds that AI adoption increases within-firm annual worker productivity growth by 2-3 percentage points. Based on this, most assume that productivity at the country level would increase by at least 1.5 percentage points in workers' average annual productivity growth over a 10-year period following the start of the productivity boom. This would represent a significant positive shock. Over the past five years, real labour productivity growth averaged 1.1% in the United States and 0.8% in the EU. A 1.5 percentage point increase would result in a tripling of productivity in the EU and a doubling in the United States (de Montpellier & Fechner, 2024).

## 4.3 AI as a Tool for Securing the Position of the EU in the Global Market

Despite promising microeconomic evidence of AI's potential to boost productivity and competitiveness, with early estimates suggesting that maintaining AI uptake among businesses in the EU could contribute an additional 600 billion EUR in gross value added to the European economy by 2030 (Strand Partners, 2024), the long-term macroeconomic outcomes remain uncertain, as they are influenced by factors such as market dynamism and functioning. While early AI adopters may experience productivity gains, these improvements may not necessarily extend to other enterprises, potentially increasing performance disparities and deepening distributional divides. AI could also have significant implications for inclusion. Furthermore, the reallocation of labour both within and across sectors could either support sustained overall growth or create a drag on progress, as extensive AI-driven automation may push a large portion of the workforce into lower-productivity roles (Filippucci et al., 2024).

Throughout history, significant technological advancements that have profoundly influenced the economy include the computer, the internet, and earlier developments like the steam engine and electricity. These are commonly known as General Purpose Technologies due to their widespread application across various technical and economic domains, fostering innovation (Bresnahan & Trajtenberg, 1995; Lipsey et al., 2005). AI can be viewed as the latest addition to this category (Varian, 2019; Agrawal et al., 2019), with the potential to drive further innovation and create a lasting positive effect on productivity, particularly when combined with other modern or emerging General Purpose Technologies like robotics and biotechnology (Cockburn et al., 2018; Yan, & Grossman, 2023). At the same time, AI has a few key features that differentiate it from previous General Purpose Technologies (Table 3) (Agrawal et al., 2023).

*Table 3: Comparing the characteristics of AI to selected previous General Purpose Technologies*

| Characteristics | Steam engine and electricity | Computers and internet | Artificial Intelligence |
|---|---|---|---|
| Main output | Energy | Calculations and information exchange | Advanced analytics (predictions, optimization) and content generation |
| Nature of tasks primarily affected | Physical | Cognitive routine and communication | Broad range of cognitive |
| Autonomy (operate independently from humans) | No | Limited | Potentially advanced |
| Capacity for self-improvement | No | Yes | Yes |
| A method of invention | No | Yes | Yes |

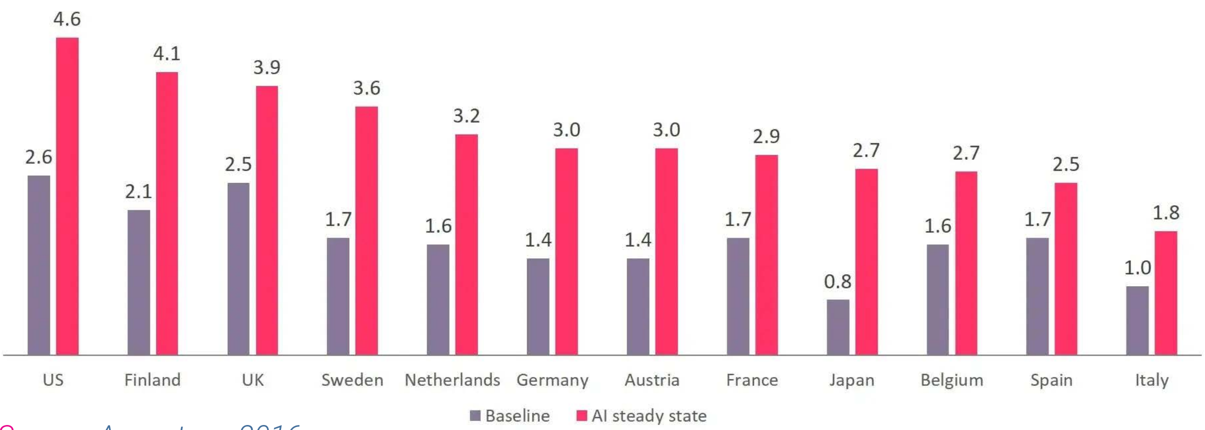
*Source: Agrawal et al., 2023; Filippucci et al., 2024; Lipsey et al., 2005.*

Previous General Purpose Technologies, such as the steam engine and electricity, primarily provided energy as their main output, which largely impacted physical tasks at the time of their invention. These technologies facilitated mechanized labour and drove industrial productivity through energy-intensive applications. In contrast, the advent of computers and the internet shifted focus toward cognitive tasks by providing intangible outputs, such as calculations and information exchange. While they affected routine cognitive tasks, these technologies had limited autonomy and no capacity for self-improvement, though they offered novel methods for invention, especially in automating certain processes or controlling physical machinery. AI, as a new General Purpose Technology, shares the cognitive task emphasis of computers and the internet but stands out by enabling more advanced analytics, such as predictions, optimizations, and content generation. AI exhibits a broader range of cognitive capabilities and can operate with a higher level of autonomy, potentially functioning independently of human intervention (Filippucci et al., 2024).

Additionally, AI has the capacity for self-improvement, allowing models to evolve without explicit programming updates and leading to increasingly complex behaviours. This capacity for self-improvement, combined with its potential autonomy, sets AI apart from previous technologies. For example, large-scale AI models can produce outputs that go beyond what is easily understood by humans, known as the “black box” phenomenon, which introduces risks of unpredictability. This unpredictability is evident in how some AI models develop unexpected capabilities, such as reasoning and language translation when scaled up in parameter size or trained on large datasets (e.g., large language models). These distinctive features – high autonomy and self-improvement – open new opportunities for innovation and welfare gains but also carry the risk of models evolving beyond human control, which calls for careful governance and regulation. Finally, the ability of AI to generate and test ideas positions it not merely as an automation technology, which primarily increases productivity in a singular, static way by enhancing existing production and processes. Instead, AI can be viewed as an invention technology, offering the potential to drive continuous productivity growth by advancing research and fostering innovation. This evolution marks a significant shift in General Purpose Technologies from energy to cognitive intelligence, raising both opportunities and challenges for the future (Filippucci et al., 2024).

Due to its significant potential, AI is discussed not only as a source of radical transformation in labour markets but also as a driver of major productivity gains that could extend beyond specific companies or sectors to benefit the entire economy. According to some experts, the technology is so revolutionary that it could lead to a positive productivity shock, potentially resulting in stronger economic growth (de Montpellier & Fechner, 2024). Despite the general belief that AI will strengthen the EU's position in the global market, there is little specific evidence on the exact potential impact of AI on economic growth in the EU, particularly due to the challenges involved in measuring this impact. However, the Accenture Institute for High Performance, in collaboration with Frontier Economics, modelled the impact of AI on 12 developed countries (9 EU countries, as well as Japan, the UK, and the US) that together generate more than 50% of the world's economic output. The analysis compared the projected size of each country's economy in 2035 under two scenarios: a baseline scenario, reflecting expected economic growth based on current conditions, and an AI scenario, which accounts for the anticipated growth after AI has been fully integrated into the economy (Figure 23). According to estimates, AI is projected to deliver the greatest economic benefits to the United States, raising its annual growth rate from 2.6% to 4.6% by 2035. In the United Kingdom, AI could increase the annual growth rate of gross value added from 2.5% to 3.9%. Japan has the potential to more than triple its gross value added growth rate by 2035, while Finland, Sweden, the Netherlands, Germany, and Austria could see their growth rates double (Accenture, 2016).

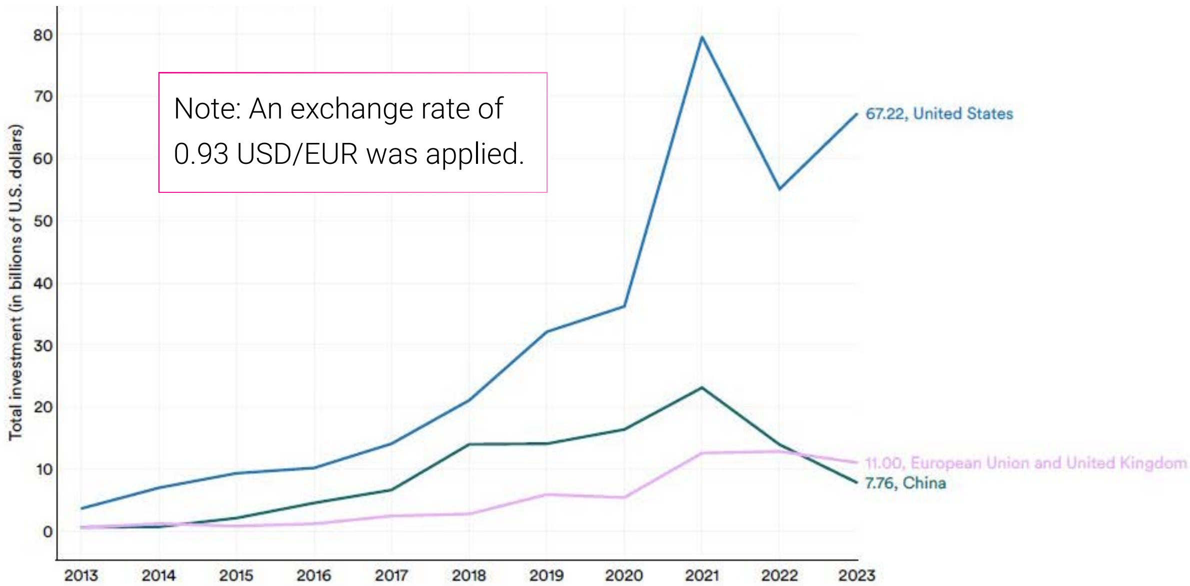
*Figure 23: Annual gross value added growth rates by 2035 (%)*




*Source: Accenture, 2016.*

The global AI market was valued at over €130 billion in 2023 and is projected to experience significant growth by 2030, reaching nearly 1.9 trillion EUR. Most of the investment in AI now comes from private sources. In 2023, the United States led in private AI investment, contributing 62.5 billion EUR, followed by China with 7.3 billion EUR (Figure 24). The EU and the United Kingdom collectively attracted 9 billion EUR in private investment in 2023. Between 2018 and the third quarter of 2023, around 32.5 billion EUR was invested in EU AI companies, compared to over 120 billion EUR in United States AI firms. Recent investments in United States companies like OpenAI and Anthropic have further widened the gap between the EU's and the US's share of private AI investment. Public investment in AI is also on the rise, with the EU's Digital Europe programme allocating a total of 2.1 billion EUR for AI from 2021 to 2027 (Madiega & Ilnicki, 2024; Maslej et al., 2024).

*Figure 24: Private investment in AI by geographic area (2013-2023)*




*Source: Maslej et al., 2024.*

At the end of the day, productivity gains at the macroeconomic level will increasingly come from the use of technology. Countries that do not produce AI technology but merely utilize it may still experience macro-level labour productivity growth. However, as seen during the productivity boom in the late 1990s, producing the technology itself will remain a critical factor. Ultimately, countries that produce AI are likely to experience more sustained productivity gains than those that only use it. The United States is expected to undergo a faster productivity surge compared to the EU. While the EU is likely to see a productivity increase, it will probably occur more slowly and remain consistently lower than the growth observed in the United States. In addition, several factors are likely to limit the impact of AI on EU productivity compared with other developed economies. First, the market for risk capital in Europe is less developed and fragmented, which hampers investment in AI innovation and the scaling of AI-driven businesses. This contrasts sharply with more unified and mature capital markets in economies like the United States, where startups and AI ventures have greater access to funding. Second, Europe's linguistic diversity presents a significant challenge for AI model development, as AI systems need to be trained across multiple languages, making the process more complex and resource-intensive compared to monolingual markets. Lastly, the EU's stricter data protection regulations, such as the GDPR, and the additional regulatory pressures anticipated under the AI Act impose tighter controls on data availability and usage. These regulatory frameworks, while ensuring greater privacy and ethical standards, could slow down AI development and deployment, further widening the productivity gap between the EU and other global leaders like the United States, where data flows more freely and regulatory environments are less restrictive (de Montpellier & Fechner, 2024)

The EU finds itself at the crossroads of AI innovation and AI regulation, navigating a delicate balance between fostering technological advancements and establishing a robust regulatory framework to secure its position in the global market. On the one hand, the EU is investing heavily in AI through programs like Horizon 2020 and the Digital Europe Program, aiming to create a thriving AI ecosystem that drives economic growth, supports startups, and enhances collaboration across its member states. On the other hand, it is setting global benchmarks for ethical AI governance through comprehensive regulations like the AI Act, which positions the EU as a global leader in promoting responsible and human-centric AI development. This dual approach is designed to ensure that AI is both an engine for innovation and a technology that aligns with European values, protecting fundamental rights and ensuring public trust. However, the path to achieving this delicate balance is fraught with challenges, as the EU's digital industry and investment levels still lag behind those of global AI leaders like the United States and China. Despite these obstacles, the EU's position at this critical intersection of innovation and regulation gives it a unique opportunity to shape the global AI landscape, provided it can successfully turn its ambitious policies into concrete action. By doing so, the EU aims not only to compete on the global stage but also to lead in setting the standards for responsible AI development worldwide (Csernatoni, 2024).

# 5 Environmental Implications of AI

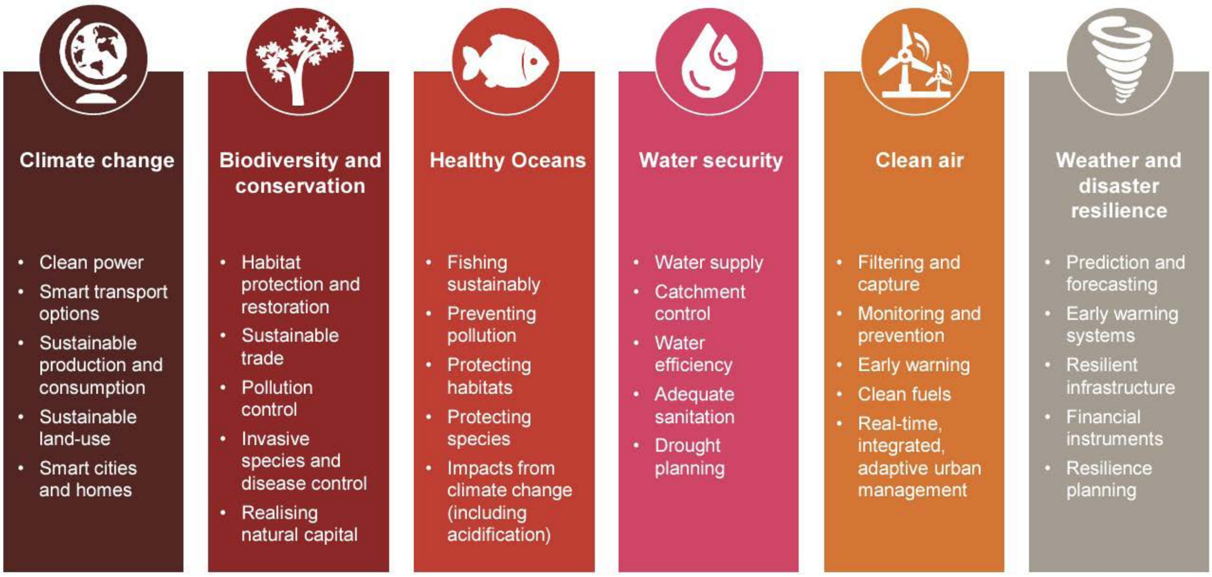
## 5.1 Supporting the Green Transition

AI is increasingly recognised as a powerful tool for tackling complex global challenges, including environmental sustainability and green transition. The rise of AI has fundamentally changed the way industry works, offering new opportunities for efficiency, resource management and decision-making. In particular, AI's ability to process vast amounts of data, recognise patterns and optimise processes makes it the ideal tool to contribute to the EU's green transition. However, the impact of AI on the environment is double-edged: while it has the potential to drive sustainable change, its own energy and material requirements also pose a challenge that needs to be addressed (PwC, 2018). As AI moves into critical sectors such as energy, agriculture and transport, there is an urgent need to align its development and deployment with broader environmental goals.

To ensure that the development of AI is in line with the EU's sustainability goals, a solid policy framework is needed. The EU AI Act introduced as the first comprehensive legal framework for AI, aims to ensure that AI systems deployed in the EU are trustworthy, safe and ethical (AI Act, 2024). While the AI Act focuses primarily on issues such as transparency, fairness and data protection, it is also in line with the EU's environmental priorities and requires AI systems to comply with sustainability goals. This legal framework is complemented by the European Green Deal (2019), the EU's strategic plan to achieve climate neutrality by 2050. The Green Deal aims to transform the European economy by reducing greenhouse gas emissions, promoting circular economies and improving biodiversity. Together, these frameworks form a basis for the integration of AI into the EU's environmental strategy and emphasise the need for a dual transition, a digital and a green one.

To fully grasp the potential of AI in the context of the green transition, it is important to consider its broad application across all sectors. Figure 25 provides a general overview of the different sectors in which AI can support the green transition and illustrates the interconnectedness of AI use in sectors like energy, agriculture, transport and industry. The figure 25 illustrates that AI is not limited to a single sector, but its role spans multiple areas, each of which contributes to the EU's climate neutrality goals (PwC, 2018). For example, the predictive capabilities of AI improve decision-making in urban planning, energy networks and transport logistics, enabling more efficient use of resources. The integration of AI into these sectors demonstrates its ability to optimise processes, reduce waste and improve the overall sustainability of the European economy.

*Figure 25: Possible use of AI in tackling environmental challenges*




*Source: PwC, 2018.*

The adoption of AI in key sectors is crucial for the realisation of the European Green Deal (2019). As shown in Figure 26, sectors such as energy, manufacturing and agriculture are early adopters of AI, using it to reduce emissions and increase efficiency. In the energy sector, AI optimises consumption and integrates renewable energy, in manufacturing it minimises resource consumption, and in agriculture, precision farming reduces environmental impact. A survey by the European Commission (2020a) found that 42% of companies are using at least one AI technology, but that barriers such as the lack of skilled labour (57%) and the cost of adoption (52%) are limiting further implementation. This underlines the need for strategic investment to overcome these barriers and fully realise the environmental potential of AI.

*Figure 26: Sectoral adoption of AI in European companies*

| Sector (Part I) | At least one AI technology | At least two AI technologies | Plans to use |
|---|---|---|---|
| Agriculture, forestry and/or fishing | 39% | 24% | 18% |
| Manufacturing | 47% | 27% | 16% |
| Construction | 36% | 23% | 16% |
| Oil and gas | 38% | 19% | 6% |
| Waste management | 31% | 21% | 27% |
| Water and electricity supply | 45% | 28% | 17% |
| Trade, retail | 38% | 22% | 20% |
| Transport | 36% | 22% | 20% |
| Food | 36% | 26% | 20% |

*Source: European Commission, 2020a.*

The transformative role of AI in the green transition is a central focus of the European Green Deal (2019) – EGD, which aims to make Europe climate neutral by 2050. Figure 27 highlights the key pillars of the EGD, showing how AI can support each of these goals. AI's ability for predictive analytics and real-time monitoring supports key EGD objectives such as sustainable energy, circular economy and pollution reduction (appliedAI, 2024). For example, AI improves the circular economy by tracking materials throughout their lifecycle, improving recycling efficiency and monitoring pollution levels to provide early warning of environmental hazards. As shown in Figure 27, policy recommendations to promote this change include building strong European AI solutions, supporting large-scale AI testbeds and creating efficient data markets for AI utilisation as part of the EGD. A societal shift in awareness of the potential of AI and targeted business support for AI applications under the EGD will further support these efforts. In addition, the development of standardised carbon assessment methods for AI applications will ensure the transparency of the carbon footprint.

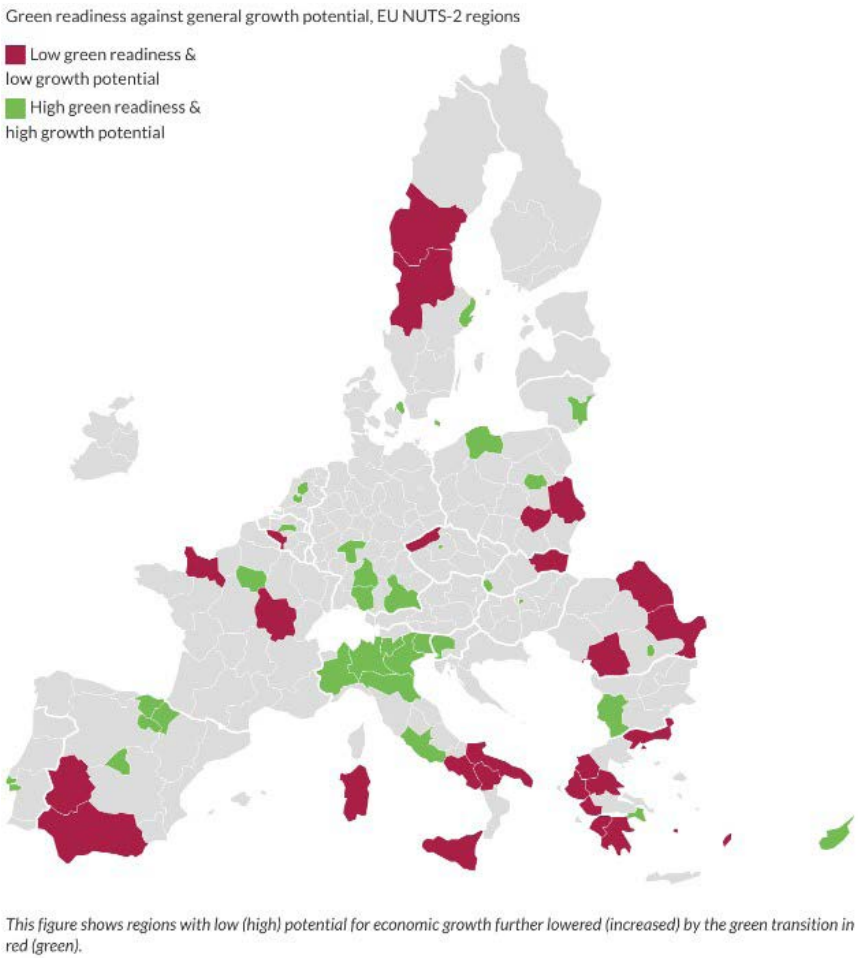
*Figure 27: Policy interventions to enable AI for EGD*

| Derived from... | Policy recommendation |
|---|---|
| Tech-nology | • Strong European solutions: Enable EU businesses to become global champions<br>• Large-scale testing beyond sandboxes: Provide experimentation envi-ronments to pilot AI applications in large-scale, real-world scenarios |
| Data | • Implement an efficient market for AI for EGD-related data in the EU |
| Skills | • Shift society's AI mindset to an opportunity-based one: Educate society in AI's potential as well as managing and understanding its risks |
| Business | • Adopt the new AI policy framework but provide administrative support for EGD cases<br>• Earmark AI funding to EGD implementation |
| Generic | • Build a pioneer coalition of EU member states that adopts AI for EGD strategies<br>• Create standardised CO2 assessment methodologies for AI appli-cations to support carbon footprint transparency |

*Source: appliedAI, 2024.*

A key challenge for the AI-driven green transition in the EU is the growing regional disparity in AI readiness and infrastructure (The Vienna Institute for International Economic Studies, 2023). Figure 28 shows that urban high-tech regions such as Germany, the Netherlands and northern Italy are well prepared to benefit from digitalisation and greening, while less developed rural areas and regions dependent on carbon-intensive industries, such as parts of eastern Europe, southern Italy, Greece, southern Spain and central Sweden, are at risk of falling behind. These regions face significant adjustment costs in switching to more environmentally friendly practises, especially in sectors such as coal mining or steel production. While these investments are necessary to reduce emissions, they may not immediately lead to an increase in economic output, which weighs on the growth potential of these areas. In the meantime, wealthier regions are better placed to benefit from the green transition due to their advanced infrastructure and technological capabilities.

*Figure 28: Strongest deviations in growth potential in relation to the green transition*




*Source: The Vienna Institute for International Economic Studies, 2023.*

The disparities in AI readiness and adoption, as well as in green sustainability infrastructure, highlight the urgent need for targeted investment to ensure that less developed regions can participate in and benefit from the green transition. Without such efforts, rural areas and regions with energy-intensive industries risk further economic decline, exacerbating inequalities in the EU. The EGD emphasises the importance of an inclusive transition, but achieving this will require significant investment in digital infrastructure and AI capabilities in lagging regions. This will help ensure that AI-driven innovation promotes both regional cohesion and sustainable development across the EU (The Green Deal, 2019).

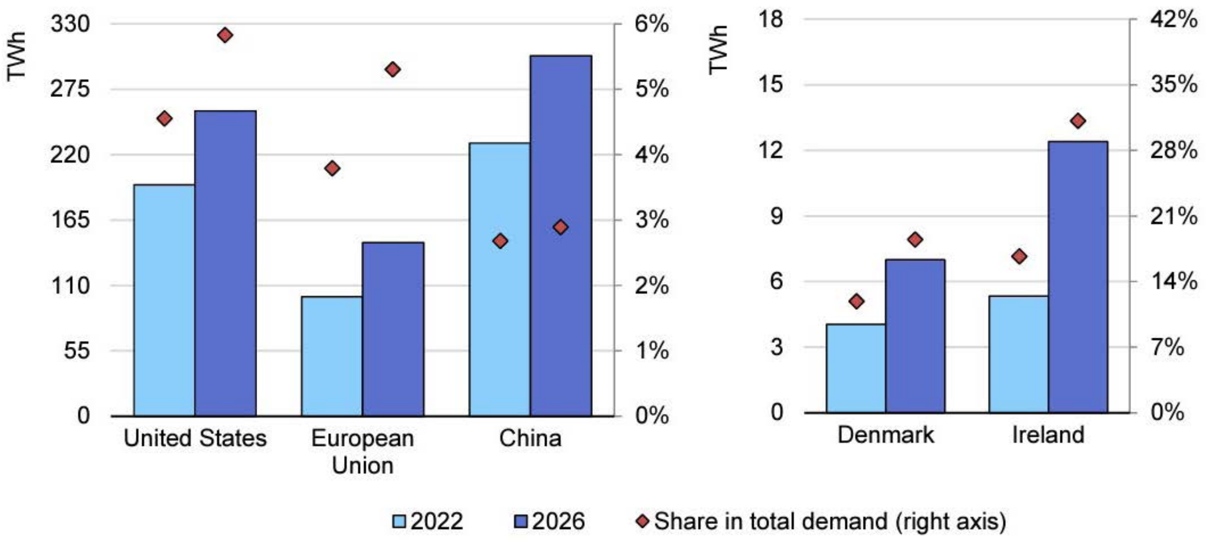
The integration of AI into the EU's green transition offers a transformative opportunity to drive sustainability in sectors such as energy, agriculture and manufacturing while aligning with the EGD. AI can optimise resource use, improve environmental monitoring and increase efficiency, supporting the EU's climate neutrality goals. However, to ensure a just transition, the EU needs to address the different challenges faced by regions dependent on fossil fuels and carbon-intensive industries. In parallel, the EU's commitment to sustainable development under the UN 2030 Agenda integrates AI into the achievement of the SDGs. With strategic investments and inclusive policies, AI can promote both environmental and social cohesion and ensure a sustainable and equitable green future for all regions of Europe (European Commission, 2024a).

## 5.2 The Carbon Footprint

As AI proliferates globally, its carbon footprint has become an important issue, particularly in relation to the energy requirements of large-scale AI models (Schwartz et al., 2020). The growing industry around AI technologies, especially machine learning models such as large-scale language models (LLMs) and neural networks, consume enormous amounts of computing power, contributing to an increase in carbon emissions (Verdecchia et al., 2023). Understanding and mitigating these environmental costs are key priorities as AI becomes more integrated into society.

Training advanced AI models requires an enormous amount of energy. This leads to a significant increase in energy consumption during both the training and deployment phases of these systems. AI systems are operated by large data centres with thousands of computing units running simultaneously to process huge data sets. These facilities consume large amounts of electricity, much of which is still generated from fossil fuels (OECD, 2022). For example, the World Economic Forum's study (2024a) on the impact of AI on the environment highlights that the energy consumption of data centres is increasing worldwide, with China, the United States and the EU having the highest consumption levels (see Figure 29). The figure shows a significant increase in the electricity consumption of data centres in selected countries such as Denmark and Ireland by 2026.

*Figure 29: Estimated electricity consumption of data centres and their share of total electricity demand in selected countries in 2022 and 2026*

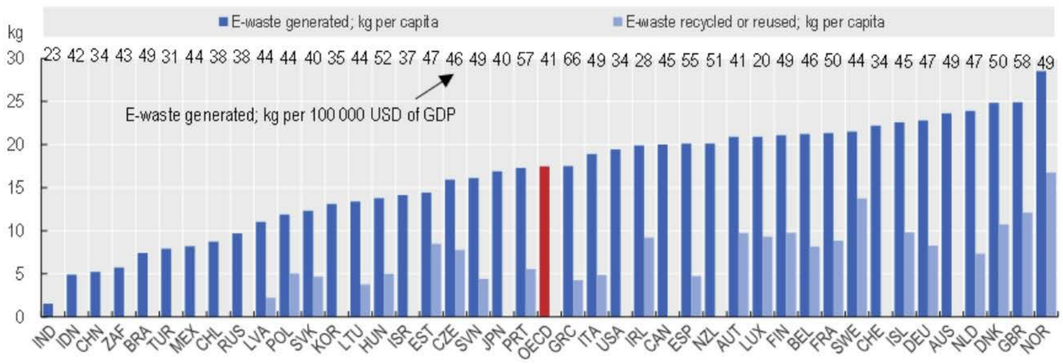



*Source: World Economic Forum, 2024a.*

The rapid expansion of cloud services and AI applications has significantly increased the demand for data storage and computing power, much of which is housed in large, energy-intensive data centres. These centres process and store the computations required for AI models and contribute significantly to the carbon emissions associated with AI technologies. (Verdecchia et al., 2023). The reliance on non-renewable energy sources such as coal and natural gas to power these data centres exacerbates their environmental impact and contributes to increased carbon emissions.

An often-overlooked consequence of this technological expansion is the increasing generation of e-waste. Figure 30 illustrates the extent of e-waste generation and recycling or reuse in different countries in 2016. Countries such as Germany, France and Sweden generate significant amounts of e-waste, including many AI-related devices such as servers, graphics processors and other computing devices. Although countries such as Norway lead the way in e-waste recycling, others, including Greece, lag behind with lower recycling rates in relation to their e-waste production. This discrepancy emphasises the environmental challenge that AI infrastructure poses, as inefficient recycling processes further increase environmental costs (OECD, 2022). To make the AI industry more sustainable, both energy consumption and e-waste generated by AI need to be tackled.

*Figure 30: E-waste generation and recycling or reuse by countries (2016)*



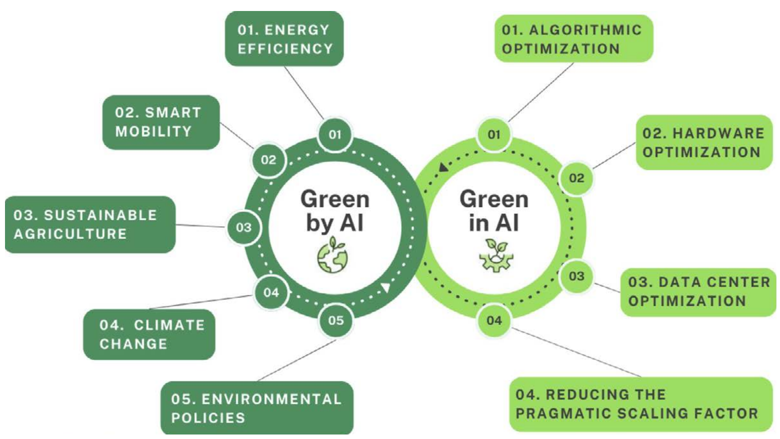

Note: This figure does not distinguish between AI and non-AI e-waste. The figure covers six waste categories: (1) temperature equipment; (2) screens, monitors; (3) lamps; (4) large equipment; (5) small equipment; and (5) small IT and telecommunication equipment. Disaggregating these categories could better estimate AI compute-related impacts.

*Source: OECD, 2022.*

Given the growing concern about the environmental footprint of AI, the concept of green AI has gained traction. Green AI emphasises the development of more energy-efficient algorithms and hardware that focus not only on performance but also on minimising environmental impact in order to reduce the carbon footprint of AI technologies (Schwartz et al., 2020).

The green AI framework distinguishes between green by AI and green in AI, as shown in Figure 31. Green by AI refers to how AI contributes to sustainability in areas such as energy efficiency, smart mobility and sustainable agriculture. Green in AI, on the other hand, focuses on making AI itself more sustainable through strategies such as algorithmic optimisation, hardware efficiency and reducing the energy consumption of data centres. This dual approach supports both AI's role in the green transition and its own environmental responsibility.

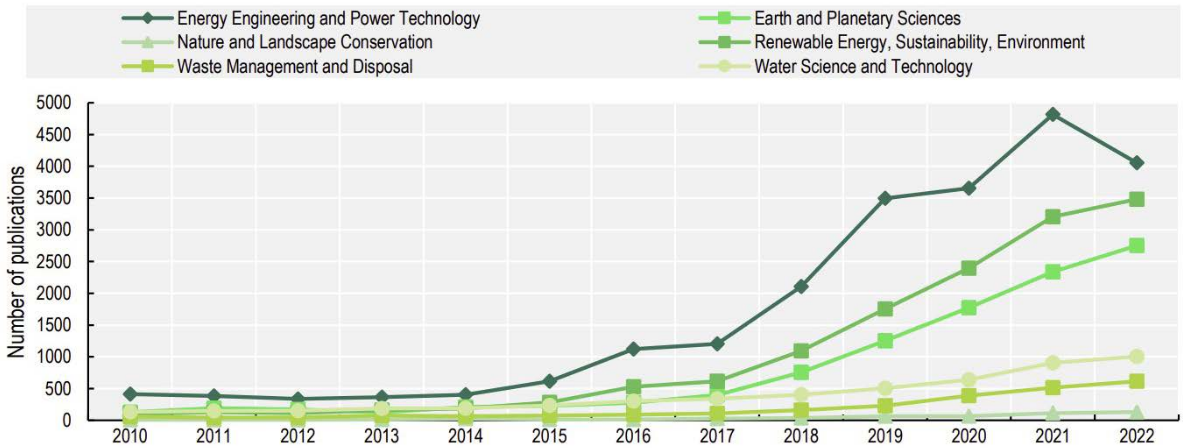
*Figure 31: Overview of green-by vs. green-in AI*




*Source: Bolón-Canedo et al., 2024.*

The growing importance of green AI is also reflected in the increasing research interest in environmentally orientated AI applications. As shown in Figure 32, the number of scientific publications on environmentally related AI topics has risen sharply, particularly in areas such as energy engineering and power technology. Since 2016, research in the fields of renewable energy, waste management and earth and planetary sciences has increased exponentially, highlighting the growing recognition of AI's potential to tackle environmental challenges. This increase in research publications shows that AI is being used not only to improve industrial efficiency but also to solve pressing global sustainability issues, emphasising the need for continued innovation in the field of green AI.

*Figure 32: Green AI publications by subject area (2010–2022)*




*Source: OECD, 2022.*

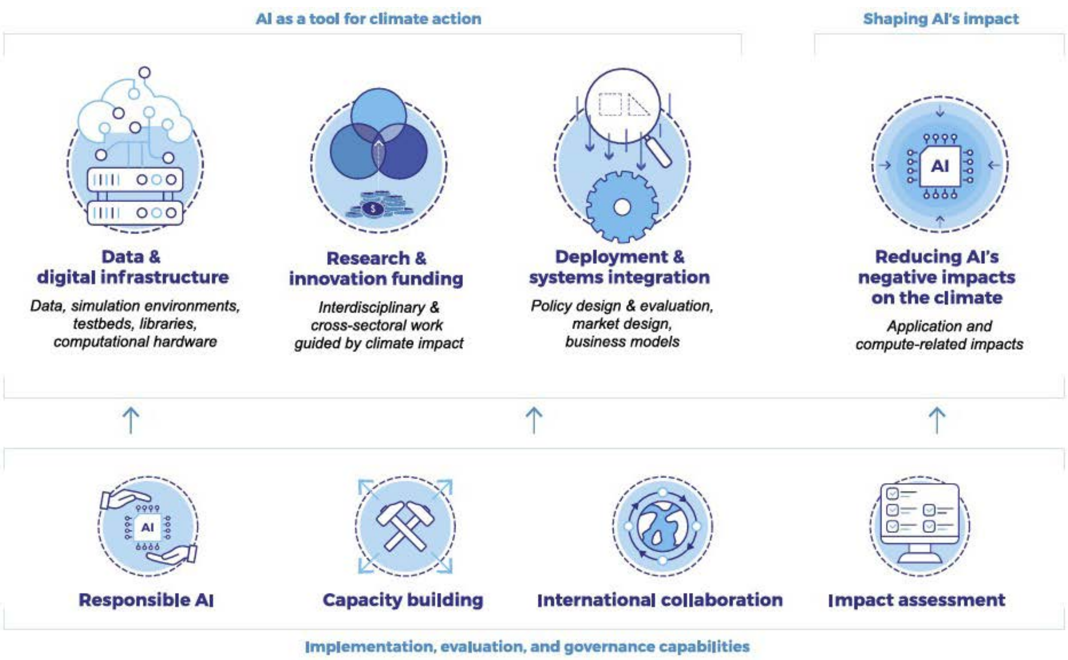
The increasing adoption and development of AI technology poses a double challenge: While AI has the potential to drive sustainability across all sectors, its carbon footprint and the generation of e-waste represent significant environmental impacts. Therefore, the rising energy consumption of data centres and the increasing amount of electronic waste need to be addressed urgently. The concept of green AI offers a promising solution by focussing on energy-efficient algorithms and infrastructures while making a positive contribution through green AI applications in multiple areas. However, for AI to truly align with global sustainability goals, concerted efforts are needed to implement green AI strategies, reduce the environmental costs of data processing and ensure responsible recycling of e-waste. Increasing scientific research on these topics shows a growing awareness, but sustained innovation and strong political support are crucial to reconcile the progress of AI with its environmental responsibility.

## 5.3 Potential Benefits and Risks for Environmental Sustainability

The use of AI technologies offers significant potential to promote environmental sustainability by enabling more efficient energy management, climate change mitigation and the sustainable utilisation of resources. However, there are also critical risks associated with the increasing adoption of these technologies, such as increased energy consumption, resource depletion and the potential for algorithmic bias. While AI can make an important contribution to achieving environmental goals, the associated risks must also be addressed in order to fully realise the benefits of AI for environmental sustainability.

Governments play a crucial role in promoting the responsible use of AI for environmental sustainability. As shown in Figure 33, investment in data infrastructure and computing resources is essential to support AI in resource management and monitoring. Research funding should focus on the development of AI innovations that improve sustainability in different areas. Policy must also address reducing the environmental impact of AI, including the energy consumption of data centres and minimising resource extraction. Ethical frameworks should ensure environmental justice, and international co-operation is needed to enhance the positive impacts of AI while mitigating environmental risks.

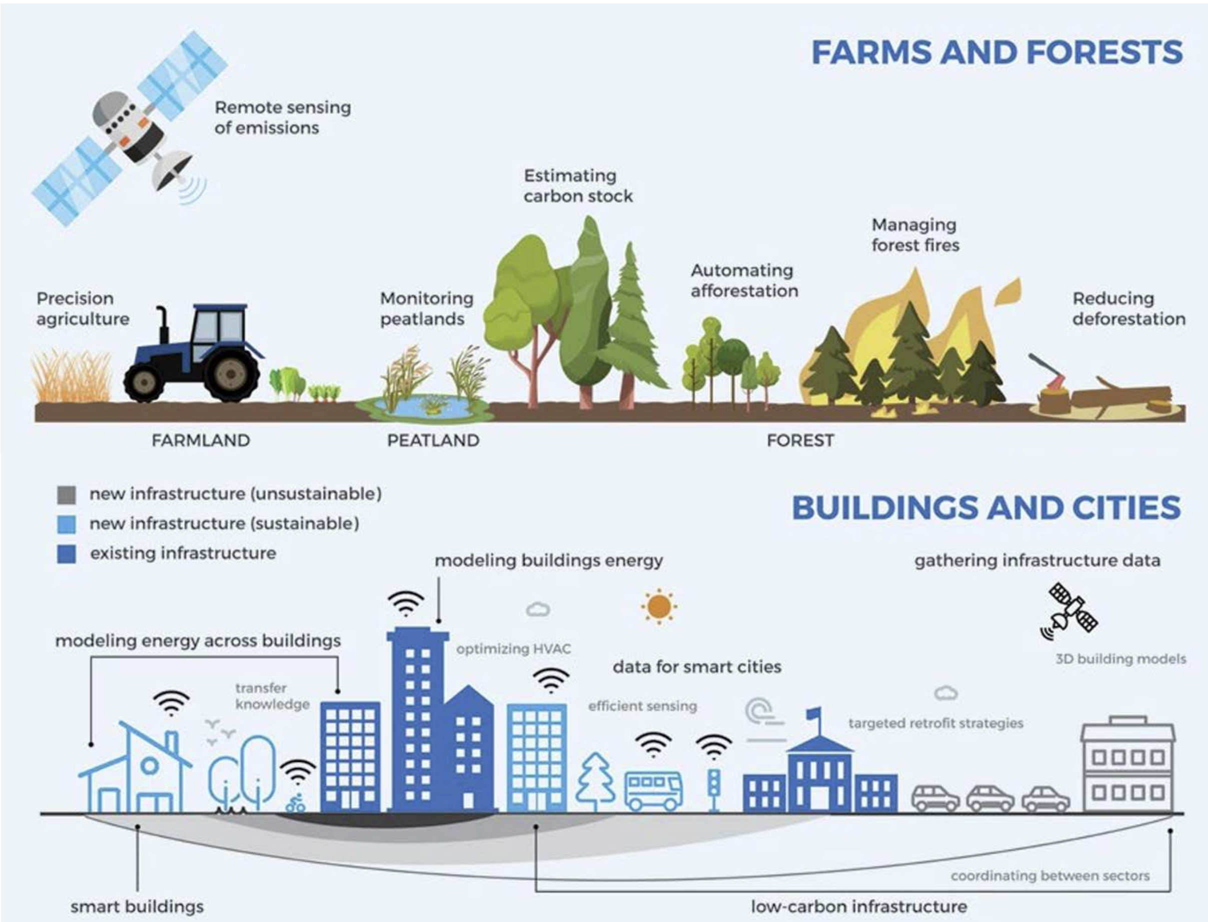
*Figure 33: Fields of action for governments in supporting the responsible use of AI in an environmental context*




*Source: Clutton-Brock et al., 2021.*

One of the most important contributions of AI to environmental sustainability is its role in mitigating climate change. AI has proven useful in modelling climate patterns and providing predictive analytics to help policy makers develop more effective climate strategies. By analysing large amounts of data, AI can improve the accuracy of climate models and help governments and organisations predict extreme weather events, assess carbon stocks and manage natural resources more sustainably (Nishant et al., 2020). For example, Figure 34 shows how AI is being used in agriculture, forestry and urban infrastructure to support sustainability efforts. In agriculture and forestry, AI tools such as remote sensing, precision agriculture and automated reforestation optimise resource use and reduce emissions. In the urban environment, AI is improving energy efficiency through smart buildings, modelling the energy consumption of buildings and data integration for smart cities, contributing to the development of low-carbon infrastructure. Overall, these applications aim to create more sustainable land and city management systems.

*Figure 34: Possible use of AI in agriculture, forestry, buildings and cities*




*Source: Clutton-Brock et al., 2021, reproduced from Kochanski et al., 2019.*

AI also plays a crucial role in the optimisation of renewable energy systems. By analysing real-time data, AI can improve the efficiency of solar and wind power generation, balance energy loads and integrate storage solutions to ensure a stable supply of renewable energy (Galaz et al., 2021). AI systems control the distribution of energy in power grids and optimise when and how energy is generated, stored and consumed. This will improve the flexibility of the energy grid and reduce energy waste, enabling greater reliance on renewables and the phasing out of fossil fuels (Clutton-Brock et al., 2021).

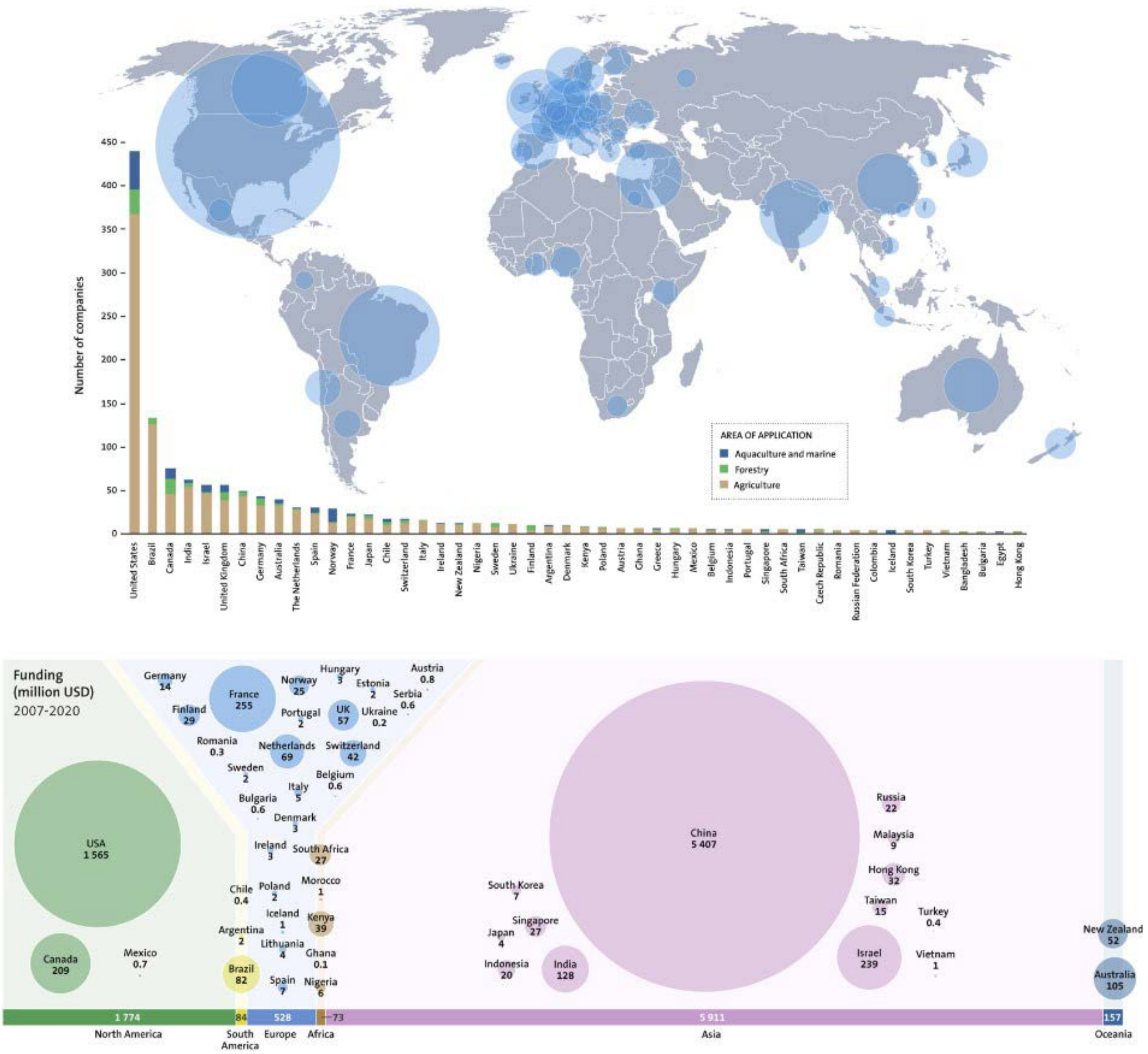
Despite its promising potential, the environmental costs of AI are not trivial. One of the main problems is the ecological toll that the production of AI-related hardware takes. The production of servers, data centres and computing devices requires rare earth metals, which are mined in environmentally harmful processes. Data centres, which are essential for the operation of AI models, are known to consume enormous amounts of electricity, much of which comes from non-renewable energy sources (Galaz et al., 2021). In addition, there are concerns about the algorithmic bias of AI systems, particularly in the context of environmental justice. AI models used to enforce environmental regulations or allocate environmentally friendly technologies could reinforce existing societal biases and disproportionately affect vulnerable populations. Communities with limited access to green technologies or energy-efficient resources may be left behind, further entrenching inequality (Nishant et al., 2020). To avoid exacerbating such inequalities, it is important to ensure that AI systems are designed and deployed equitably.

Another critical point is the regulatory gaps in AI development in many countries, especially in the EU. The rapid growth of AI has outpaced the introduction of comprehensive environmental regulations, which has led to negative environmental effects. Without strong regulatory frameworks, AI development could contribute to unsustainable practises, from excessive energy consumption in data centres to inappropriate treatment of electronic waste (European Environment Agency, 2023).

The White Paper on AI (European Commission, 2020b) emphasises that while AI can make a significant contribution to environmental sustainability, it must be regulated to minimise its environmental impact. Key areas include responsible data use, energy-efficient computational methods and strict ethical oversight. The paper argues for a harmonised regulatory framework across Europe to ensure that AI development is in line with sustainability goals, particularly in the context of the European Green Deal. This balance between innovation and regulation is essential to prevent AI from exacerbating environmental challenges while maximising its potential for positive environmental impact.

Furthermore, as Figure 35 shows, within Europe, countries such as France, Germany and the Netherlands are leading AI-driven sustainability efforts, which is reflected in both funding and the number of companies involved. However, the concentration of resources in these countries emphasises the regional differences within Europe itself. In many Eastern European countries and smaller EU member states, there is significantly less investment in AI technologies for sustainability and fewer companies involved. This imbalance raises concerns about unequal access to AI innovations and their potential benefits, which could hinder wider, continent-wide progress towards environmental goals.

*Figure 35: Global distribution of AI technologies and investments in farming, forestry and the marine/aquaculture sectors*




*Source: Galaz et al., 2021.*

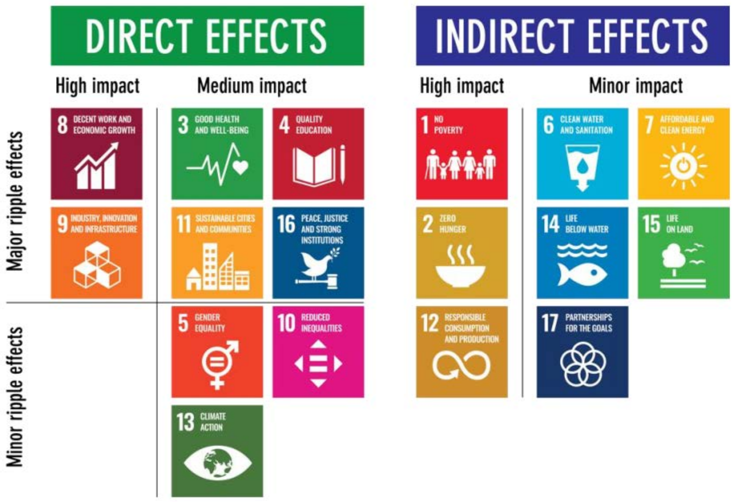
AI's potential for environmental sustainability is undeniable, but both its benefits and risks must be considered. On the positive side, AI offers numerous advantages, such as optimising resource management, mitigating climate change, and improving renewable energy systems. However, significant risks must also be addressed, including environmental concerns such as energy consumption, carbon footprint, resource depletion, and the risk of distortion. To achieve the optimal balance between these outcomes and promote environmental sustainability, the following key aspects should be considered:

- **Concerns over AI's energy use:** The rapid expansion of AI infrastructure, especially energy-intensive data centres, raises sustainability concerns. AI's reliance on non-renewable energy sources contributes significantly to carbon emissions, increasing the technology's carbon footprint and sustainability challenges.
- **Environmental impact of AI hardware:** The production of AI hardware, including servers and computing devices, incurs high environmental costs, particularly through the extraction of rare earth minerals. This often-overlooked resource depletion adds complexity to AI's sustainability.
- **Global inequalities in AI access:** Developed regions benefit from AI, but many underdeveloped areas lack the infrastructure and resources to leverage its environmental benefits. This creates a significant gap in investment and access, limiting AI's broader societal impact.
- **Proactive measures for responsible AI development:** To ensure AI supports environmental sustainability, governments and organisations must prioritise ethical frameworks, adopt energy-efficient AI technologies, and enforce recycling regulations for AI-related e-waste.
- **Policy recommendations for sustainability:** Governments should introduce standards for energy-efficient AI models and data centres and provide incentives for using renewable energy in AI operations. Policymakers must also ensure equitable access to AI across regions to prevent disparities.
- **Creating a favourable policy environment:** Policymakers must address AI's environmental challenges throughout its lifecycle, from development to deployment, to make AI a key driver of sustainable progress while protecting the planet.
- **Fostering global collaboration:** Global collaboration and public-private partnerships are vital for deploying green AI in key sectors like energy, agriculture, and transport. Collaborative innovation can bridge regulatory gaps and align AI with sustainability goals.
- **Addressing regional inequalities:** Policymakers must address disparities in AI-driven sustainability benefits between regions. Strategic investments and inclusive policies are needed to ensure AI's environmental and socio-economic advantages are shared equitably.
- **Long-term vision for AI and sustainability:** A long-term approach should integrate digital, green, and smart transitions, ensuring AI aligns with climate neutrality and sustainability goals. Governments should also expand AI's role in environmental monitoring and encourage energy-efficient innovation.

# 6 Policy Recommendations

AI is associated with various socioeconomic and environmental impacts, prompting significant attention toward analyzing and evaluating these effects. As AI continues to evolve, its influence on areas such as healthcare, education, labour markets, economic development, resource consumption, and the natural environment becomes more pronounced. To understand the broader impact of AI on social, economic, and environmental areas, policymakers should look beyond immediate benefits and consider the long-term effects of AI, ensuring it aligns with sustainable development and the corresponding SDGs. Specifically, by evaluating both the direct and indirect effects of AI, it is essential to assess how its implementation influences interconnected systems across various sectors. This approach helps policymakers identify potential ripple effects, such as unintended consequences in employment, inequality, or environmental degradation and emphasizes the importance of responsible innovation that aligns with long-term socioeconomic and environmental sustainability. There are several social, economic, and environmental impacts of AI that can be operationalized and discussed through the lens of the SDGs, where the SDGs are grouped based on the level of impact, whether AI primarily affects efforts to achieve the SDGs directly or indirectly, and whether the goals have clear and significant ripple effects (Sætra, 2021) (Figure 36).

*Figure 36: Categorisation of the SDGs based on the AI effects*




*Source: Sætra, 2021.*

AI plays a pivotal role in advancing the United Nations SDGs across various dimensions: social, economic, and environmental. AI technologies influence these areas both directly and indirectly, creating ripple effects that extend across multiple goals. In the following, there is a brief presentation of the social, economic, and environmental impacts of AI, focusing on how AI shapes progress toward these global objectives (Sætra, 2021):

- **Social impact of AI:** AI significantly affects social SDGs by directly influencing goals related to health, education, and justice. AI technologies are transforming SDG 3 (Good Health and Well-being) by enhancing diagnostics, personalized medicine, and health system management, leading to improved healthcare outcomes. In education (SDG 4), AI-based learning systems and remote education platforms provide more inclusive and equitable learning opportunities. AI also indirectly impacts SDG 16 (Peace, Justice, and Strong Institutions) by promoting transparency in governance and improving justice systems through predictive algorithms. The ripple effects of these direct impacts on social goals are substantial, as better health and education systems contribute to a more equitable society (SDG 10) and gender equality (SDG 5).
- **Economic impact of AI:** The economic impact of AI is one of its most significant contributions, as seen in SDG 8 (Decent Work and Economic Growth) and SDG 9 (Industry, Innovation, and Infrastructure). AI enhances productivity, fosters innovation, and supports the creation of new industries. These developments drive economic growth and job creation, though automation can also lead to job displacement. AI's major ripple effects extend from economic growth, influencing SDG 1 (No Poverty) and SDG 2 (Zero Hunger) by indirectly reducing poverty and improving food security through more efficient agricultural systems. Additionally, AI contributes to infrastructure development and technological innovation, which further sustains economic growth and improves living standards across various regions.
- **Environmental impact of AI:** AI's impact on environmental SDGs is highlighted by its contributions to SDG 13 (Climate Action), SDG 12 (Responsible Consumption and Production), and ecosystem management goals like SDG 14 (Life Below Water) and SDG 15 (Life on Land). AI technologies enable more efficient use of resources, optimize energy systems, and monitor environmental changes, aiding efforts to combat climate change. While AI contributes to environmental sustainability, its overall impact is more limited compared to social and economic goals. Moreover, AI's development itself poses environmental challenges, particularly due to the energy consumption of large-scale data processing. Therefore, while AI aids in monitoring and managing environmental resources, the need to minimize its own carbon footprint is a growing concern.

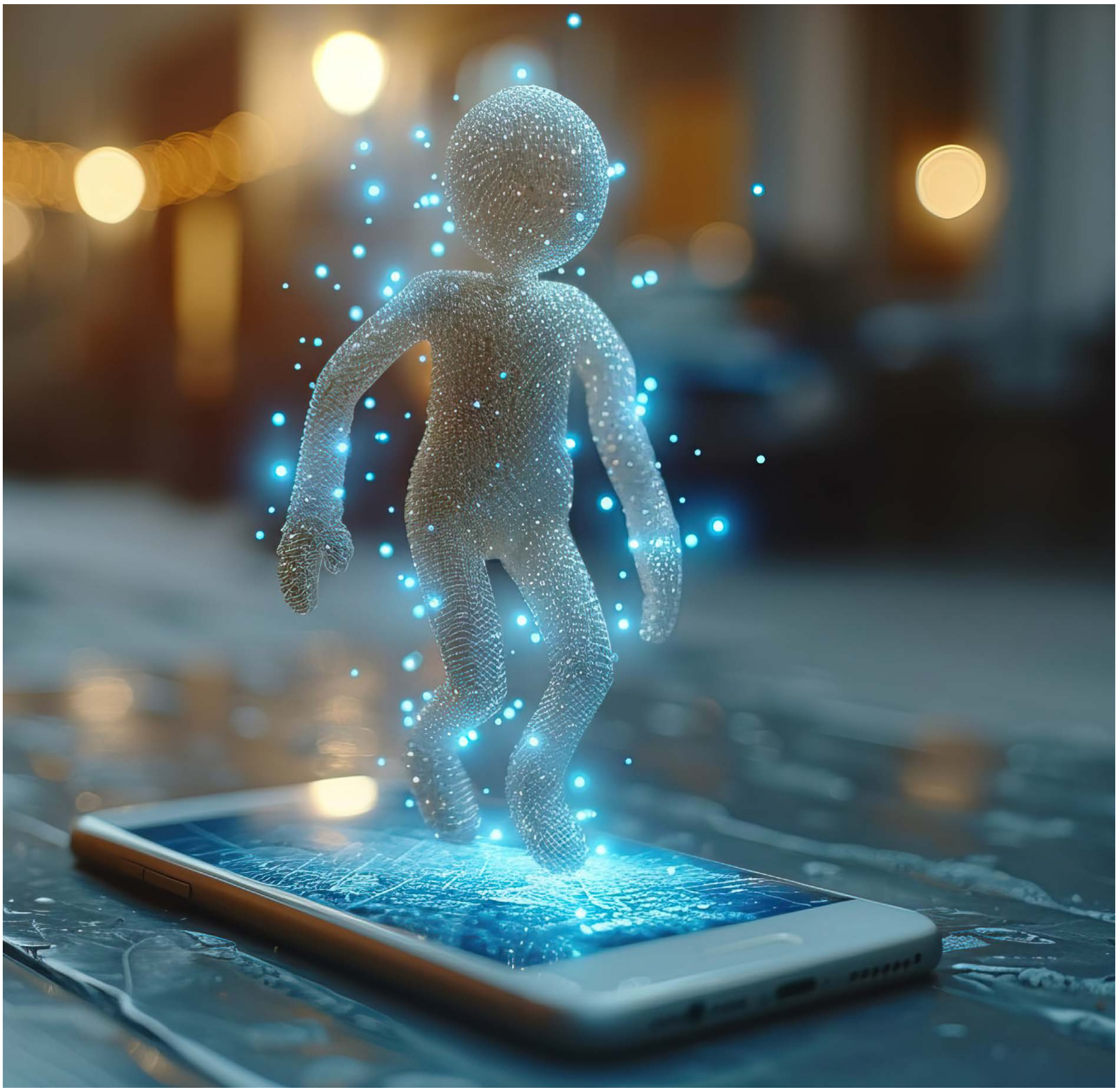
In order to address the multifaceted societal, economic, and environmental impacts of AI within the EU, this publication proposes a set of recommendations aimed at fostering the responsible development, deployment, and regulation of AI systems. These recommendations align with the broader goals of ensuring equity, sustainability, and inclusivity, while safeguarding fundamental human rights and promoting competitive economic growth in the digital era. Below are 11 key recommendations for policymakers, practitioners, and other stakeholders (adapted from European Commission, 2019b, 47-49):

1. **Promote AI ecosystems through sectoral multi-stakeholder collaboration:** Collaboration between civil society, industry, research institutions, and public authorities is essential to building vibrant AI ecosystems. Establishing multi-stakeholder alliances at a sectoral level will allow for a more nuanced understanding of the specific challenges and opportunities that AI presents across different industries, such as healthcare, education, and energy. By fostering partnerships, stakeholders can ensure that AI systems address the needs of their specific sectors, promoting more efficient public services and innovative industrial processes.
2. **Ensure a single European market for AI:** To maintain Europe's global competitiveness, it is crucial to secure a unified single European market for AI. This requires harmonising legislation across EU member states, avoiding market fragmentation, and maintaining high standards for protecting individual rights and freedoms. A cohesive regulatory framework will promote the free flow of AI technologies and services, allowing businesses to scale AI solutions across borders and compete globally while upholding the EU's ethical standards.
3. **Invest in AI-driven economic growth and competitiveness:** AI holds the potential to drive significant economic growth by enhancing productivity and creating new markets. Strategic investment in AI research and development, along with supporting infrastructure such as high-performance computing and data-sharing platforms, will position the EU as a leader in the global AI race. Policymakers should also promote public-private partnerships that align AI innovation with the EU's long-term economic and environmental goals, ensuring that AI drives sustainable growth.
4. **Empower and protect individuals and society:** Public trust in AI is essential for its successful adoption. Empowering citizens with the knowledge and skills to engage with AI systems is critical. Policymakers should focus on raising awareness of AI's capabilities and limitations, ensuring that individuals are equipped with the skills needed to navigate a workforce increasingly shaped by AI technologies. In addition, robust safeguards must be put in place to protect citizens from the potential risks AI poses, including biased decision-making and infringements on privacy.
5. **Foster a European data economy:** Data is the cornerstone of AI innovation. The EU must create a vibrant data economy by establishing frameworks that facilitate data access, sharing, and interoperability, while also maintaining strict privacy and security standards. This requires investment in infrastructure to support large-scale data collection and processing. By fostering a European data economy, the EU can enable the development of trustworthy AI systems and unlock the potential for data-driven innovation in key sectors such as healthcare, transport, and energy.

6. **Address global inequalities in AI access and use:** AI presents an opportunity to drive socioeconomic equality, but disparities in access to AI technologies could exacerbate existing inequalities. Policymakers must focus on closing the digital divide between more developed and underdeveloped regions, ensuring that all areas of the EU can benefit from AI's potential. This includes investing in infrastructure and education in less developed regions, fostering AI-related skills, and ensuring that AI systems are designed and deployed inclusively so that they benefit all citizens equally.
7. **Exploit the public sector's role as an AI leader:** The public sector has a pivotal role in demonstrating the responsible use of AI. Governments can lead by example by integrating AI into public service delivery, thereby improving efficiency and transparency while maintaining strong safeguards for fundamental rights. Through innovation-driven public procurement, the public sector can foster the development of AI solutions that address societal challenges and improve public services. This leadership is critical for building trust in AI and ensuring that it serves the public good.
8. **Strengthen and coordinate AI research capabilities:** The EU must strengthen its AI research capabilities to remain at the forefront of technological innovation. By creating additional centres of excellence in AI and increasing collaboration between academia, industry, and public institutions, the EU can build a unified AI research landscape. A coordinated research agenda that focuses on solving grand challenges, such as climate change, healthcare, and sustainable agriculture, will ensure that AI development aligns with Europe's strategic priorities and long-term goals.
9. **Adopt risk-based governance for AI:** Effective AI governance requires a nuanced, risk-based approach that is both proportionate and adaptable to emerging challenges. The EU should ensure that regulatory frameworks are updated to address the specific risks associated with AI, including biased algorithms, privacy concerns, and the potential for job displacement. A comprehensive review of existing laws is necessary to ensure they remain fit for purpose in an AI-driven world, while additional regulatory mechanisms should be introduced where needed to protect individuals and encourage innovation.
10. **Nurture AI education and continuous learning:** The EU needs a deep and broad skills base to support its AI ambitions. Educational systems should integrate AI literacy into primary, secondary, and tertiary curricula, while also ensuring lifelong learning opportunities that allow individuals to upskill and reskill for the AI-driven economy. Fostering interdisciplinary and multidisciplinary learning will help individuals better understand the societal and environmental implications of AI, ensuring that Europe's workforce remains competitive and adaptable in the face of rapid technological change.
11. **Stimulate a robust investment environment:** Increased investment in AI is critical to ensuring the EU's leadership in AI development and deployment. The Horizon Europe and

Digital Europe programmes provide a solid foundation, but further efforts are needed to stimulate private-sector investment. Policymakers should focus on creating a favourable environment for AI innovation by building trust among stakeholders, fostering sectoral alliances, and ensuring that investments are directed towards responsible AI that aligns with the EU's ethical principles and environmental goals.

By implementing these recommendations, the EU can lay the foundation for a future where AI drives sustainable growth, enhances societal well-being, protects the environment, and establishes Europe as a global leader in ethical AI. A comprehensive and balanced approach to AI governance, investment, and development will ensure that AI not only contributes to a greener, fairer, and more prosperous Europe, but also safeguards fundamental rights and promotes inclusivity. By fostering innovation and collaboration across sectors, the EU can ensure that AI benefits all citizens while maintaining its competitive edge on the global stage.



# 7 Conclusion

As AI continues to evolve from a niche technology into a transformative force, its potential to reshape entire sectors—ranging from healthcare and education to labour markets and environmental sustainability—offers the EU an unprecedented opportunity to lead in innovation and technological advancement on a global scale. However, this transformation must be handled with caution to ensure it aligns with the EU's commitment to safeguarding human rights, promoting social equity, and ensuring sustainable development.

The publication's analysis of AI's societal, economic, and environmental dimensions illustrates the dual nature of this technology. On the one hand, AI can drive significant progress in improving healthcare outcomes, enhancing productivity across industries, and contributing to environmental goals such as energy efficiency and resource optimisation. On the other hand, AI presents complex ethical, economic, and ecological challenges. These include potential job displacement, growing economic inequality, environmental degradation from AI infrastructure, and the risk of perpetuating or exacerbating societal biases. The EU, through regulatory initiatives like the AI Act, is uniquely positioned to address these challenges. The AI Act's emphasis on categorising AI systems by risk, enhancing transparency, and ensuring accountability provides a solid framework for ensuring that AI development remains aligned with Europe's core values of inclusivity, fairness, and respect for fundamental rights. This regulatory approach is pivotal in shaping the global discourse on AI ethics and governance, ensuring that AI benefits society while minimising its risks.

For AI to reach its full potential in the EU, several key actions must be prioritised. First, policymakers and industry leaders need to focus on creating robust regulatory frameworks that address transparency, fairness, and accountability. Ensuring that AI systems are explainable, free from bias, and used responsibly is essential for fostering public trust and preventing misuse. Second, targeted investments in workforce reskilling and education are crucial to prepare workers for the new labour landscape that AI is creating. This will help mitigate the risks of job displacement while promoting social and economic inclusion. Additionally, as AI's environmental footprint grows, the concept of "green AI" must be at the forefront of AI development. Reducing the carbon footprint of data centres, improving energy efficiency in AI operations, and advancing research into sustainable AI technologies are essential to ensure that AI supports, rather than hinders, the EU's environmental and climate neutrality goals. Collaboration across all sectors—public, private, and academic—will be key to the successful deployment of AI in Europe. This cooperation will ensure that AI continues to drive innovation while adhering to the EU's strong ethical and environmental standards. By embracing these guiding principles, the EU can not only lead the world in AI governance but also ensure that this powerful technology is used to advance human well-being, social equity, and environmental sustainability.

In conclusion, AI presents both extraordinary opportunities and significant challenges. If guided by the principles of transparency, fairness, accountability, and sustainability, AI has the potential to greatly enhance the quality of life across Europe and globally. By implementing the recommendations outlined in this publication, the EU can harness AI's transformative power while ensuring that its deployment aligns with its core values, ultimately leading to a future where AI serves the collective good, promotes equity, and supports long-term social, economic and environmental progress.

# 8 References


1. Accenture. (2016). Artificial Intelligence Poised to Double Annual Economic Growth Rate in 12 Developed Economies and Boost Labor Productivity by up to 40 Percent by 2035, According to New Research by Accenture. Dublin: Accenture. https://newsroom.accenture.com/news/2016/artificial-intelligence-poised-to-double-annual-economic-growth-rate-in-12-developed-economies-and-boost-labor-productivity-by-up-to-40-percent-by-2035-according-to-new-research-by-accenture
2. Agrawal, A. K., Gans, J. S., & Goldfarb, A. (2023). Similarities and differences in the Adoption of General Purpose Technologies (No. w30976). National Bureau of Economic Research. https://doi.org/10.3386/w30976
3. Agrawal, A., Gans, J., & Goldfarb, A. (2019). Economic policy for artificial intelligence. Innovation policy and the economy, 19(1), 139-159. https://doi.org/10.1086/699935
4. Amazon Team. (2024). AI adoption forecast to unleash €600 billion growth in Europe's economy. Seatlle: Amazon Team. https://www.aboutamazon.eu/news/job-creation-and-investment/ai-adoption-forecast-to-unleash-600-billion-growth-in-europes-economy
5. AppliedAI, Initiative for applied artificial intelligence. (2024). How to leverage AI to support the European Green Deal. Munich: appliedAI Initiative. https://www.appliedai.de/assets/files/How-to-leverage-AI-to-support-the-European-Green-Deal_2024-03-01-105553_hbio.pdf
6. Aristovnik, A., Umek, L., & Ravšelj, D. (2023a). Artificial Intelligence in Public Administration: A Bibliometric Review in Comparative Perspective. In Conference on Information Technology and its Applications (pp. 126-140). Cham: Springer Nature Switzerland.
7. Aristovnik, A., Umek, L., Brezovar, N., Keržič, D., & Ravšelj, D. (2023b). The Role of ChatGPT in Higher Education: Some Reflections from Public Administration Students. In International Conference on Technology in Education (pp. 254-263). Singapore: Springer Nature Singapore.
8. Babšek, M., Murko, E., & Aristovnik, A. (2024). Adoption of artificial intelligence in public administration: Bibliometric study and systematisation of practical implementation. In EGPA 2024 Conference, September 3–6, 2024, Athens, Greece (pp. 1–19). Athens: EGPA.
9. Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. Journal of AI, 7(1), 52-62. https://dx.doi.org/10.2139/ssrn.4337484
10. Banfa, A. (2024). Introduction to Artificial Intelligence (AI). Gistrup: River Publishers. https://doi.org/10.1201/9781003499527
11. Bankhwal, M., Bisht, A., Chui, M., Roberts, R., & van Heteren, A. (2024). AI for social good: Improving lives and protecting the planet. McKinsey Digital. https://www.mckinsey.com/capabilities/quantumblack/our-insights/ai-for-social-good
12. Bhatia, A. (2024). AI's Exponential Journey: Milestones to AGI and Beyond. https://www.linkedin.com/pulse/ais-exponential-journey-milestones-agi-beyond-ashish-bhatia-o5lle/

13. Bresnahan, T. F., & Trajtenberg, M. (1995). General purpose technologies 'Engines of growth'?. Journal of econometrics, 65(1), 83-108. https://doi.org/10.1016/0304-4076(94)01598-T
14. Bolón-Canedo, V., Morán-Fernández, L., Cancela, B., & Alonso-Betanzos, A. (2024). A review of green artificial intelligence: Towards a more sustainable future. Neurocomputing, 128096. https://doi.org/10.1016/j.neucom.2024.128096
15. Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work (No. w31161). Cambridge: National Bureau of Economic Research. https://www.nber.org/system/files/working_papers/w31161/w31161.pdf
16. Busch, P. A., & Henriksen, H. Z. (2018). Digital discretion: A systematic literature review of ICT and street-level discretion. Information Polity, 23(1), 3–28. http://dx.doi.org/10.3233/IP-170050
17. Cazzaniga, M., Jaumotte, F., Li, L., Melina, G., Panton, A. J., Pizzinelli, C., Rockall, E., & Tavares, M. M. (2024). Gen-AI: Artificial intelligence and the future of work. International Monetary Fund. https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2024/01/14/Gen-AI-Artificial-Intelligence-and-the-Future-of-Work-542379
18. Cengage Group. (2024). 2024 Graduate Employability Report: Preparing students for the GenAI-driven workplace. Boston: Cengage Group. https://cengage.widen.net/s/bmjxxjx9mm/cg-2024-employability-survey-report
19. Clutton-Brock, P., Rolnick, D., Donti, P. L., & Kaack, L. (2021). Climate change and AI. recommendations for government action. GPAI, Climate Change AI, Centre for AI & Climate. https://www.gpai.ai/projects/climate-change-and-ai.pdf
20. Cockburn, I. M., Henderson, R., & Stern, S. (2018). The impact of artificial intelligence on innovation (Vol. 24449). Cambridge, MA, USA: National bureau of economic research. https://doi.org/10.7208/9780226613475-006
21. Csernatoni, R. (2024). Charting the Geopolitics and European Governance of Artificial Intelligence. Brussels: Carnegie Europe. https://carnegieendowment.org/research/2024/03/charting-the-geopolitics-and-european-governance-of-artificial-intelligence?lang=en&center=europe
22. de Montpellier, C., & Fechner, I. (2024). AI productivity gains may be smaller than you're expecting. Amsterdam: ING Think. https://think.ing.com/articles/macro-level-productivity-gains-ai-coming-artificial-intelligence-the-effect-smaller/
23. Deloitte. (2023). Generative AI and the future of work. London: Deloitte. https://www2.deloitte.com/content/dam/Deloitte/us/Documents/consulting/us-ai-institute-generative-ai-and-the-future-of-work.pdf
24. DPO Centre. (2024). Compliance with the AI Act Part 1: Timeline and important deadlines. https://www.dpocentre.com/compliance-with-the-ai-act-what-you-need-to-know/
25. Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., ... & Williams, M. D. (2019). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002
26. European Commission. (2024a). Green transition. https://reform-support.ec.europa.eu/what-we-do/green-transition_en
27. European Commission. (2024b). Public Sector Tech Watch latest dataset of selected cases. European Commission [Dataset]. https://joinup.ec.europa.eu/collection/public-sector-tech-watch/data-download
28. European Commission. (2021). Coordinated plan on artificial intelligence 2021 review. Brussels: European Commission. https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review

29. European Commission. (2020a). European enterprise survey on the use of technologies based on artificial intelligence. Brussels: European Commission. https://digital-strategy.ec.europa.eu/en/library/european-enterprise-survey-use-technologies-based-artificial-intelligence

30. European Commission. (2020b). White Paper On Artificial Intelligence - A European approach to excellence and trust. COM (2020) 65. Brussels: European Commission. https://commission.europa.eu/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en

31. European Commission. (2019a). Ethics guidelines for trustworthy AI. Independent High-Level Expert Group on Artificial Intelligence set up by the European Commission. Brussels: European Commission. https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

32. European Commission. (2019b). Policy and investment recommendations for trustworthy AI. Brussels: European Commission. https://www.europarl.europa.eu/cmsdata/196378/AI%20HLEG_Policy%20and%20Investment%20Recommendations.pdf

33. European Commission. (2018). Communication from the Commission: Artificial Intelligence for Europe. Brussels: European Commission. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN

34. European Commission. (2012). Charter of Fundamental Rights of the European Union (2012/C 326/02). Official Journal of the European Union, C 326, 391–407. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A12012P%2FTXT

35. European Environment Agency. (2023). Accelerating the circular economy in Europe. State and outlook 2024. Copenhagen: European Environment Agency. https://www.eea.europa.eu/publications/accelerating-the-circular-economy

36. European Green Deal. (2019). COM/2019/640. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2019%3A640%3AFIN

37. Eurostat. (2024). Use of artificial intelligence in enterprises. Luxembourg: Eurostat. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Use_of_artificial_intelligence_in_enterprises

38. Filippucci, F., Gal, P., Jona-Lasinio, C., Leandro, A., & Nicoletti, G. (2024). The impact of Artificial Intelligence on productivity, distribution and growth: Key mechanisms, initial evidence and policy challenges. OECD Artificial Intelligence Papers. No. 15, OECD Publishing, Paris. https://doi.org/10.1787/8d900037-en

39. Floridi, L., & Cowls, J. (2022). A unified framework of five principles for AI in society. Machine learning and the city: Applications in architecture and urban design, 535-545. http://dx.doi.org/10.1162/99608f92.8cd550d1

40. Galaz, V., Centeno, M. A., Callahan, P. W., Causevic, A., Patterson, T., Brass, I., ... & Levy, K. (2021). Artificial intelligence, systemic risks, and sustainability. Technology in Society, 67, 101741. https://doi.org/10.1016/j.techsoc.2021.101741

41. Guliyev, H. (2023). Artificial intelligence and unemployment in high-tech developed countries: new insights from dynamic panel data model. Research in Globalization, 7, 100140. https://doi.org/10.1016/j.resglo.2023.100140

42. Heikkilä, M. (2024). The AI Act is done. Here's what will (and won't) change. MIT Technology Review. https://www.technologyreview.com/2024/03/19/1089919/the-ai-act-is-done-heres-what-will-and-wont-change/

43. Illing, S. (2023). Stuart Russell wrote the textbook on AI safety. He explains how to keep it from spiralling out of control. Vox. https://www.vox.com/the-gray-area/23873348/stuart-russell-artificial-intelligence-chatgpt-the-gray-area

44. Jimeno Serrano, J. F. (2024). Artificial intelligence and employment: maximising the benefits and minimising the negatives. Madrid: Banco de España. https://www.bde.es/wbe/en/noticias-eventos/blog/inteligencia-artificial-y-empleo-como-explotar-sus-potenciales-beneficios-y-evitar-sus-efectos-adversos.html

45. Kochanski, K., Rolnick, D., Donti, P., & Kaack, L. (2019). Climate change+ AI: Tackling climate change with machine learning. In AGU Fall Meeting Abstracts (Vol. 2019, pp. GC33A-04). http://dx.doi.org/10.48550/arXiv.1906.05433

46. Lane, M., Williams, M., & Broecke, S. (2023). The impact of AI on the workplace: Main findings from the OECD AI surveys of employers and workers. OECD Social, Employment and Migration Working Papers, No. 288, OECD Publishing, Paris. https://doi.org/10.1787/ea0a0fe1-en

47. Lipsey, R. G., Carlaw, K. I., & Bekar, C. T. (2005). Economic transformations: general purpose technologies and long-term economic growth. Oup Oxford.

48. Madiega, T., & Ilnicki, R. (2024). AI investment: EU and global indicators. Strasbourg: European Parliament. https://www.europarl.europa.eu/RegData/etudes/ATAG/2024/760392/EPRS_ATA(2024)760392_EN.pdf

49. Maslej, N., Fattorini, L., Perrault, R., Parli, V., Reuel, A., Brynjolfsson, E., Etchemendy, J., Ligett, K., Lyons, T., Manyika, J., Niebles, J. C., Shoham, Y., Wald, R., & Clark, J. (2024). The AI Index 2024 Annual Report. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University. https://aiindex.stanford.edu/wp-content/uploads/2024/05/HAI_AI-Index-Report-2024.pdf

50. McKinsey & Company. (2024). The state of AI in early 2024: Gen AI adoption spikes and starts to generate value. New York: McKinsey & Company. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai#/

51. McKinsey Global Institute. (2019). Tech for good: Using technology to smooth disruption and improve well-being. New York: McKinsey & Company. https://www.mckinsey.com

52. Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). AI and education: A guidance for policymakers. Paris: Unesco Publishing. https://gcedclearinghouse.org/resources/ai-and-education-guidance-policy-makers

53. Murko, E., Ravšelj, D., Umek, L., & Aristovnik, A. (2023). Artificial intelligence in public administration: Evidence from bibliometric analysis. In IIAS-SEAPP Doha Conference 2023: Developmental States and Professionalization of Public Administration and Public Policy, February 6–9, 2023, Doha, Qatar (pp. 1–20). Doha: IIAS.

54. Nishant, R., Kennedy, M., & Corbett, J. (2020). Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda. International Journal of Information Management, 53, 102104. https://doi.org/10.1016/j.ijinfomgt.2020.102104

55. OECD. (2024a). Explanatory memorandum on the updated OECD definition of an AI system. OECD Artificial Intelligence Papers, No. 8. Paris: OECD Publishing. https://doi.org/10.1787/623da898-en

56. OECD. (2024b). The potential impact of artificial intelligence on equity and inclusion in education. OECD Artificial Intelligence Papers, No. 23. Paris: OECD Publishing. https://www.oecd-ilibrary.org/science-and-technology/the-potential-impact-of-artificial-intelligence-on-equity-and-inclusion-in-education_15df715b-en

57. OECD. (2022). Measuring the environmental impacts of artificial intelligence compute and applications. The AI footprint. Paris: OECD Publishing. https://www.oecd.org/en/publications/measuring-the-environmental-impacts-of-artificial-intelligence-compute-and-applications_7babf571-en.html

58. OECD. (2019). Artificial Intelligence in Society. Paris: OECD Publishing. https://doi.org/10.1787/eedfee77-en.

59. OpenMinded. (2022). Of Legal Tangles and Synthetic Datasets Part 3: GDPR and Synthesis. https://blog.openmined.org/of-legal-tangles-and-synthetic-datasets-gdpr-and-synthesis/

60. PwC. (2018). Fourth Industrial Revolution for the Earth. Harnessing Artificial Intelligence for the Earth. PwC. https://www.pwc.com/gx/en/sustainability/assets/ai-for-the-earth-jan-2018.pdf

61. Ravšelj, D., & Aristovnik, A. (2024). The perception of public administration students on ChatGPT implications for the labour market: A comparison across academic disciplines. In EGPA 2024 Conference, September 3–6, 2024, Athens, Greece (pp. 1–10). Athens: EGPA.

62. Ravšelj, D., & Babšek, M. (2024). The AI Act Balancing Innovation and Regulation in Welfare: A Case Study of the e-Welfare System in Slovenia. In A. Aristovnik, D. Ravšelj, & N. Tomaževič (Eds.), The Regulation of Artificial Intelligence: Shaping Governance and Power Relations for a Liberal Future. European Liberal Forum.

63. Regulation (EU) 2016/679 of the European Parliament and of the Council, General Data Protection Regulation. (2016). https://eur-lex.europa.eu/eli/reg/2016/679/oj

64. Regulation (EU) 2022/868 of the European Parliament and of the Council, Data Governance Act. (2022). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0868

65. Regulation (EU) 2022/1925 of the European Parliament and of the Council, Digital Markets Act. (2022).https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv%3AOJ.L_.2022.265.01.0001.01.ENG&toc=OJ%3AL%3A2022%3A265%3ATOC

66. Regulation (EU) 2022/2065 of the European Parliament and of the Council, Digital Services Act. (2022). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065

67. Regulation (EU) 2024/1689 of the European Parliament and of the Council, AI Act. (2024). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206

68. Righi, R., Leon, C. P., Cardona, M., Garrido, J. S., Papazoglou, M., Samoili, S., & Baillet, M. V. P. (2022). AI watch index 2021 (No. JRC128744). Luxembourg: Joint Research Centre. https://dx.doi.org/10.2760/921564

69. Sætra, H. S. (2021). AI in context and the sustainable development goals: Factoring in the unsustainability of the sociotechnical system. Sustainability, 13(4), 1738. https://doi.org/10.3390/su13041738

70. Samuel, A. L. (1960). Some moral and technical consequences of automation—a refutation. Science, 132(3429), 741-742. https://doi.org/10.1126/science.132.3429.741

71. Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. Communications of the ACM, 63(12), 54-63. https://doi.org/10.1145/3381831

72. Shuliak, M. (2024). AI in healthcare: Examples, use cases, and benefits. Acropolium. https://acropolium.com/blog/ai-in-healthcare-examples-use-cases-and-benefits/

73. Sidley. (2024). EU Formally Adopts World's First AI Law. https://datamatters.sidley.com/2024/03/21/eu-formally-adopts-worlds-first-ai-law/

74. Siebel, T. M. (2019). Digital transformation: survive and thrive in an era of mass extinction. New York: RosettaBooks.

75. SoftProdigy. (n.d.). How is AI chatbot development transforming the education sector? SoftProdigy. https://softprodigy.com/how-is-ai-chatbot-development-transforming-the-education/

76. Sok, S., & Heng, K. (2023). ChatGPT for education and research: A review of benefits and risks. Cambodian Journal of Educational Research, 3(1), 110-121. http://dx.doi.org/10.2139/ssrn.4378735

77. Stanford Institute. (2024). Artificial Intelligence Index Report 2024. Stanford Institute for Human-Centered Artificial Intelligence (HAI). https://aiindex.stanford.edu/report/

78. Strand Partners. (2024). Unlocking Europe's AI Potential in the Digital Decade. London: Strand Partners. https://www.unlockingeuropesaipotential.com/

79. Suresh, H., & Guttag, J. (2021). A framework for understanding sources of harm throughout the machine learning life cycle. In Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization (pp. 1-9). https://doi.org/10.48550/arXiv.1901.10002

80. Verdecchia, R., Sallou, J., & Cruz, L. (2023). A systematic review of Green AI. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 13(4), e1507. https://doi.org/10.1002/widm.1507

81. Tai, M. C. T. (2020). The impact of artificial intelligence on human society and bioethics. Tzu Chi Medical Journal, 32(4), 339-343. https://doi.org/10.4103/tcmj.tcmj_71_20

82. Takyar, A. (n.d.). AI in education: Use cases, benefits, solution and implementation. LeewayHertz. https://www.leewayhertz.com/ai-use-cases-in-education/

83. The Vienna Institute for International Economic Studies. (2023). Green and digital transformation will widen the gap between poor and rich regions in the EU. https://wiiw.ac.at/green-and-digital-transformation-will-widen-the-gap-between-poor-and-rich-regions-in-the-eu-n-597.html

84. Trail. (n.d.). EU AI Act: How risk is classified. Trail. https://www.trail-ml.com/blog/eu-ai-act-how-risk-is-classified

85. Tomaževič, N., Murko, E., & Aristovnik, A. (2024). Organizational Enablers of Artificial Intelligence Adoption in Public Institutions: A Systematic Literature Review. Cent. Eur. Pub. Admin. Rev., 22, 109.

86. Ulnicane, I. (2022). Artificial Intelligence in the European Union: Policy, ethics and regulation. In The Routledge handbook of European integrations. London and New York: Taylor & Francis.

87. United Nations. (2023). The Sustainable Development Goals Report 2023: Special edition. New York: United Nations. https://unstats.un.org/sdgs/report/2023/The-Sustainable-Development-Goals-Report-2023.pdf

88. United Nations. (n.d.). The 17 goals. United Nations. https://sdgs.un.org/goals

89. Varian, H. R. (2018). Artificial intelligence, economics, and industrial organization (Vol. 24839). Cambridge, MA, USA: National Bureau of Economic Research. https://doi.org/10.7208/9780226613475-018

90. World Economic Forum. (2024a). AI and energy: Will AI help reduce emissions or increase demand? Here's what to know. https://www.weforum.org/agenda/2024/07/generative-ai-energy-emissions/

91. World Economic Forum. (2024b). AI for Impact: The Role of Artificial Intelligence in Social Innovation. Geneva: World Economic Forum. https://www.weforum.org/publications/ai-for-impact-artificial-intelligence-in-social-innovation/

92. World Economic Forum. (2023). The Future of Jobs Report 2023. Geneva: World Economic Forum. https://www.weforum.org/publications/the-future-of-jobs-report-2023/

93. Yan, L., & Grossman, G. M. (2023). Robots and AI: A new economic era. Taylor & Francis. https://library.oapen.org/handle/20.500.12657/55791

94. Yang, G. Z., Bellingham, J., Dupont, P. E., Fischer, P., Floridi, L., Full, R., ... & Wood, R. (2018). The grand challenges of science robotics. Science robotics, 3(14), eaar7650. https://doi.org/10.1126/scirobotics.aar7650

freiheit.org