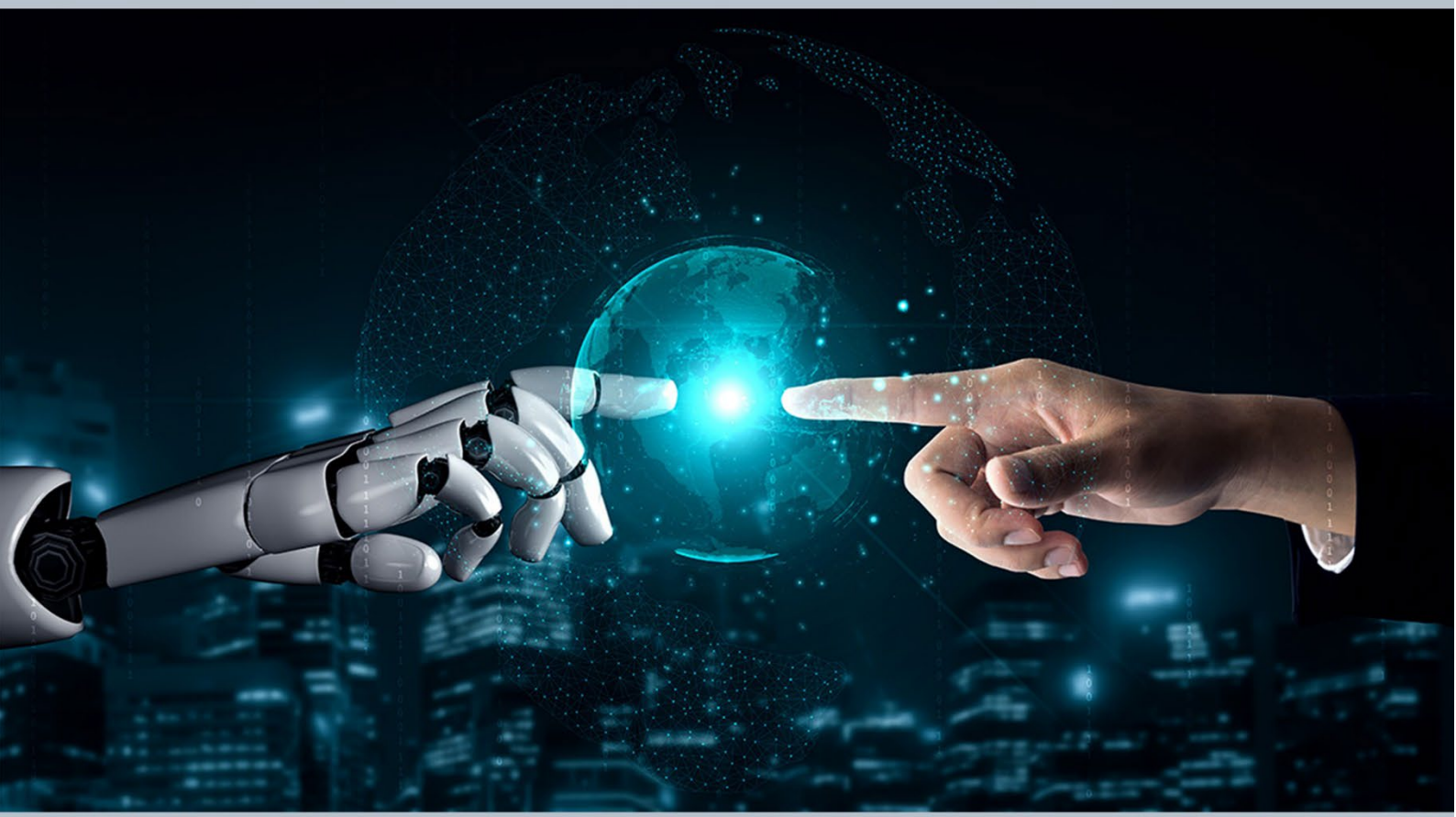
refrAIme

Co-funded by the European Union

Reinforcing Equality and Fundamental Rights in an Artificial Intelligence-Maintained Environment

Background study on the fundamental rights implications of the use of AI



February 2025

CSD Center for the Study of Democracy

Centre for European Constitutional Law

L-Università ta' Malta Faculty of Laws | Department of European & Comparative Law

Estonian Human Rights Centre

European Center for Not-for-Profit Law

Preface

This document is Deliverable 2.1 "Background study on the fundamental rights implications of the use of AI" of the EU-funded project *Reinforcing Equality and Fundamental Rights in an Artificial Intelligence-Maintained Environment* (Project: 101141304 — REFRAIME — CERV-2023-CHAR-LITI). Its main purpose is to record the legal and policy framework, as well as practical applications of Artificial Intelligence (AI), at the time of writing, in order to serve as a resource for the project's capacity building and awareness-raising activities, and serve as the basis of Deliverable 2.2 "Policy brief on the fundamental rights implications in the use of AI".

The Study was drafted by the Centre for European Constitutional Law – Themistokles and Dimitris Tsatsos Foundation, based on contributions from the project partners: the Center for the Study of Democracy (Coordinator), the European Center for Not-For-Profit Law Stichting, the Estonian Human Rights Centre, and the University of Malta. It is based on extensive desk and qualitative field research with the participation of fifty (50) stakeholders, including legal professionals, AI developers, and policymakers. In addition, three (3) interviews were conducted with representatives of international and European organisations with a relevant remit – including fundamental rights, AI, democracy and the rule of law – who wished to provide information anonymously and without mention of their organisation.

Deliverable information

Deliverable No	2.1
Deliverable title	Background study on the fundamental rights implications of the use of AI
Draft No	3 (05.02.2025)
Deliverable due date	28.02.2025
Date of delivery	04.03.2025
Author(s)	Zoe Kasapi (CECL)
Contributor(s)	Dimitar Markov, Tatyana Novossiolova (CSD) Liina Laanpere, Egert Rünne (EHRC), Oleksandr Pastukhov (UM), Karolina Iwańska (ECNL)

Co-funded by the European Union

Table of contents

Co-funded by the European Union

Co-funded by
the European Union

List of Abbreviations

AI	Artificial Intelligence
AI HLEG	High Level Expert Group on AI
CAHAI	Ad hoc Committee on Artificial Intelligence
CAI	Committee on Artificial Intelligence
CEB	United Nations System Chief Executives Board for Coordination
CFREU	Charter of Fundamental Rights of the European Union (Charter)
CoE	Council of Europe
CJEU	Court of Justice of the European Union
DESI	Digital Economy and Society Index
DPA	Data Protection Authority
EC	European Commission
ECHR	European Convention on Human Rights
ENISA	European Union Agency for Cybersecurity
ETSI	European Telecommunications Standards Institute
EU	European Union
FRA	European Union Agency for Fundamental Rights
FRIA	Fundamental Rights Impact Assessment
GA	General Assembly
GDPR	General Data Protection Regulation
GPAI	Global Partnership on AI
HLCM	UN High-level Committee on Management
HLCP	UN High-Level Committee on Programmes
IASC	Inter-agency Standing Committee
IAWG-AI	UN Inter-Agency Working Group on Artificial Intelligence
IOT	Internet of Things
LLM	Large Language Model
NLP	Natural Language Processing
OECD	Organisation for Economic Cooperation and Development
PACE	Parliamentary Assembly of the Council of Europe
SMEs	Small and Medium Enterprises
SCS	Social Credit System

Co-funded by
the European Union

SDGs	Sustainable Development Goals
UN	United Nations
UNCRC	United Nations Convention on the Rights of the Child
USA	United States of America

Co-funded by
the European Union

List of Figures

List of Tables

Co-funded by the European Union

Introduction

"If we are to harness the benefits of artificial intelligence and address the risks, we must all work together - governments, industry, academia and civil society - to develop the frameworks and systems that enable responsible innovation. […] We must seize the moment, in partnership, to deliver on the promise of technological advances and harness them for the common good."

UN Secretary-General António Guterres, AI for Good Global Summit, Geneva, 2019

***

The rapid advancement and integration of artificial intelligence (AI) technologies have brought transformative changes across various sectors, offering considerable opportunities while also raising significant social and ethical challenges. In recent years, the global "AI summer" has witnessed a surge in AI policymaking and governance initiatives, reflecting both the promise and the risks associated with this technology. Governments, international organizations, and diverse stakeholders have been mobilizing efforts to ensure AI's benefits are realized while addressing its potential to undermine fundamental rights. Canada became the first country [1] to announce a national AI strategy in 2017. At the international level, the G7 adopted a "common vision for the future of artificial intelligence" in 2018, [2]committing to promote human-centric AI through appropriate technical, ethical and technologically neutral approaches that prioritise privacy, cybersecurity and data protection, while also fostering investment and innovation in AI technologies. The UN has adopted ethical guidelines [3] and, recently, a General Assembly (GA) Resolution. [4]

Following closely, the European Union (EU) published an EU Strategy for Artificial Intelligence for Europe and a Coordinated Plan on AI in 2018 (addressed in more detail in the relevant section on EU approaches). These strategic documents outline the main tensions evident in most AI-related policy and governance frameworks: leading the technological race while mitigating the associated risks to fundamental rights and the social fabric. Aiming at *"ensuring that AI works for people and is a force for good*

[1] The Pan-Canadian AI Strategy, accessible at https://cifar.ca/ai/.
[2] The Charlevoix common vision for the future of artificial intelligence, accessible at https://www.international.gc.ca/world-monde/international_relations-relations_internationales/g7/documents/2018-06-09-artificial-intelligence-artificielle.aspx?lang=eng.
[3] *Recommendation on the Ethics of Artificial Intelligence*, UNESCO, SHS/BIO/PI/2021/1, Published on 23 November 2021, accessible at https://unesdoc.unesco.org/ark:/48223/pf0000381137.
[4] General Assembly (GA) Resolution A/78/L.49 of 11 March 2024 *"Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development"*, accessible at https://undocs.org/Home/Mobile?FinalSymbol=A%2F78%2FL.49&Language=E&DeviceType=Desktop&LangRequested=False

Co-funded by the European Union

in society", the EU adopted in the summer of 2024 Regulation (EU) 2024/1689, also known as the AI Act. As the first legal text to be binding and directly applicable in for all EU Member States (MS), the AI Act adopts a risk-based approach to AI systems, aiming to prohibit certain AI practices which are deemed incompatible with human dignity and fundamental rights and regulate high-risk applications of this technology. Moreover, the European Commission (EC) has signed on behalf of the EU the Council of Europe (CoE) Framework Convention on Artificial Intelligence and human rights, democracy and the rule of law, the first-ever international legally binding treaty, open for signature since September 2024 to CoE Member States and beyond (addressed in more detail in the relevant section on international approaches).

The present study focuses on the impact of AI technologies on fundamental rights, as enshrined in the Charter of Fundamental Rights of the European Union (CFREU). It draws on research performed by the REFRAIME partners, including desk and field research involving legal professionals and AI developers in five EU Member States – Bulgaria, Estonia, Greece, the Netherlands, and Malta, as well as international actors. It aims to identify broader ethical and fundamental rights concerns through concrete use cases from the partner countries and beyond, in order to provide practical examples and serve as a repository of resources that will feed into the project's capacity building and awareness-raising activities. At the same time, the Background Study can be viewed as a useful tool for EU and national policymakers, highlighting current and future risk areas, and assisting in the operationalisation of the ethical and fundamental rights principles enshrined in the applicable legal and policy texts.

Risks

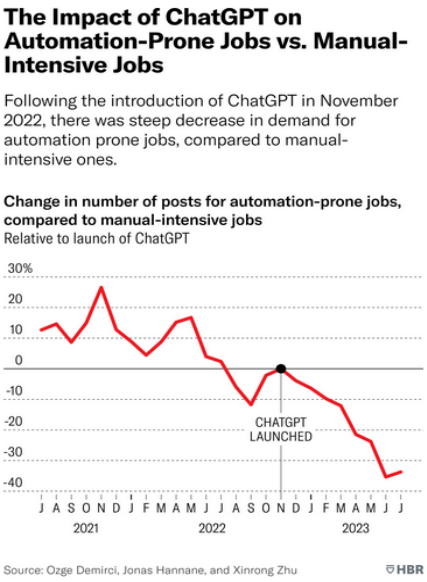
The above-mentioned developments are a clear indication of the profound and far-reaching impact of AI on a global scale. AI is reshaping how both states and private actors operate, driving innovation, and unlocking efficiencies across sectors such as healthcare, education, finance, manufacturing, and transportation. At the individual level, AI has deeply embedded itself into our daily routines, transforming how we interact and engage with reality. It influences our entertainment, our consumer habits, our communications, and various other facets of our private lives. The risks spawning from this entanglement are already felt. More than that, the ubiquitous presence of this technology threatens to leave behind anyone who doesn't adapt, disproportionately affecting socially vulnerable people. Keeping in mind the definition of 'risk', adopted in the AI Act, as the combination of the probability of an occurrence of harm and the severity of that harm, certain critical areas emerge where urgent intervention is needed to preserve fundamental rights and core EU values. Nevertheless, it will

Co-funded by
the European Union

become evident over the course of this study that a close assessment of the impacts of AI is warranted in relation to most, if not all, sectors.

Automation and AI-driven jobs can have a significant impact on the **labour market**, causing unprecedented short-term job replacement rates. [5] Moreover, AI-enabled surveillance in the workplace, as well as protest management, can severely curtail **workers' rights**, especially in relation to collective bargaining and fair and just working conditions. Another element to consider here is the protection of the so-called personality rights. 2023 was marked by intense collective action, led by artists' unions across the United States of America (USA), to negotiate rules regarding the use of digital replicas and synthetic performers. However, some commentators argue that the deal that was struck does not address the issue of **training data and the use of copyrighted material**, owned by the studios, which can be used to potentially create unlimited synthetic performers. [6]

Figure 1: The Impact of ChatGPT on automation-Prone Jobs



Source: Harvard Business Review

Privacy and personal data protection can be at risk when it comes to the use of vast sets of training data in machine learning. This has raised concerns and prompted action from national Data Protection Authorities (DPAs). However, training data sets are far the only threat these rights face. Evidence shows that authoritatarian regimes purposely collect p ersonal data without a clear justification to use to exert control over people and the civil society. For example, China's Social Credit System (SCS) uses AI for **facial recognition, behaviour tracking and analysis, and citizen grading, in combination with mass surveillance** through its estimated 700 million CCTV cameras, to enforce specific, state-sanctioned behaviours. [7] Some have argued that similar technologies are currently being used in

[5] *Research: How Gen AI Is Already Impacting the Labor Market*, Ozge Demirci, Jonas Hannane, and Xinrong Zhu, Harvard Business Review, November 11, 2024, Accessible at https://hbr.org/2024/11/research-how-gen-ai-is-already-impacting-the-labor-market.

[6] *The SAG-AFTRA Strike is Over, But the AI Fight in Hollywood is Just Beginning*, Matt Scherer, Center for Democracy and Technology, January 4, 2024, Accessible at https://cdt.org/insights/the-sag-aftra-strike-is-over-but-the-ai-fight-in-hollywood-is-just-beginning/.

[7] China's Social Credit System: A Challenge to Human Rights, Quan Van Nguyen, Sébastien Lafrance, Cu Thanh Vu, Law, State and Telecommunications Review, Volume 15(2), 2023, EISSN 1984-8161, Accessible at

Co-funded by the European Union

European countries, too, pointing to the French police and their use of **predictive policing** software. This includes facial recognition software which is deployed for risk assessment and monitoring of delinquency. Critics note that the use of this software raises serious concerns over privacy and data protection; that it reinforces biases resulting in **discrimination** and that it can lead to power abuse which undermines **personal liberty**. [8]

While these applications of AI raise serious concerns for the rights of all persons, they can be particularly problematic when applied to vulnerable populations. For example, invasive surveillance systems and decision-making algorithms applied in the refugee camps of Samos, Greece, have garnered widespread criticism for **criminalising refugees and asylum seekers** and violating their **privacy** and **dignity**, as well as their rights to **asylum and non-refoulement**. [9] As EU funds are increasingly diverted toward border control and management, [10] pilot projects like iBorderCtrl, [11] which used a virtual border guard to ask questions to third-country nationals crossing EU borders and analysed their micro-gestures for lie-detecting, have been lambasted by the civil society as "dystopian". [12] iBorderCtrl has been scrutinised by the Court of Justice of the EU (CJEU) for lack of transparency in the use of risky technologies [13] and is unlikely to be officially deployed. However, it has set a dangerous precedent for the adoption of similar technologies in this field. [14] Women's rights are also under threat. Various applications of AI technologies disproportionately impact them compounding on pre-existing societal norms and structures that disfavour them. Technology-facilitated gender-based violence has exploded in recent years, through generative AI and deepfake technologies. Video and image-based abuse is a

https://www.researchgate.net/publication/381756396_China's_Social_Credit_System_A_Challenge_to_Human_Rights.

[8] *Predictive Policing in France: Against opacity and discrimination, the need for a ban*, La Quadrature du Net, 18 January 2024, Accessible at https://www.laquadrature.net/en/2024/01/18/predictive-policing-in-france-against-opacity-and-discrimination-the-need-for-a-ban/.

[9] *Automation and Surveillance in Fortress Europe: The Digital Walls of Fortress Europe - Part 3*, Kostas Zafeiropoulos, Ioanna Louloudi, Nikos Morfonios, Mediterranean Institute for Investigative Reporting (MIIR), 19/5/2022, accessible at https://miir.gr/en/automation-and-surveillance-in-fortress-europe/.

[10] See *Funds for Fortress Europe: spending by Frontex and eu-LISA*, Statewatch, 28 January 2022, accessible at https://www.statewatch.org/analyses/2022/funds-for-fortress-europe-spending-by-frontex-and-eu-lisa/.

[11] Implemented between 2016-2019, funded by the European Union's Horizon 2020 research and innovation programme, and piloted in Greece, Latvia and Hungary. For more information, see https://web.archive.org/web/20190731191128/https://www.iborderctrl.eu/.

[12] https://www.reuters.com/article/technology/eus-lie-detecting-virtual-border-guards-face-court-scrutiny-idUSL8N2KB2GT/.

[13] https://edri.org/our-work/european-court-supports-transparency-in-risky-eu-border-tech-experiments/.

[14] Petra Molnar, Technological *Testing Grounds: Migration Management Experiments and Reflections from the Ground Up*, European Digital Rights and the Refugee Law Lab, 2020, accessible at https://edri.org/wp-content/uploads/2020/11/Technological-Testing-Grounds.pdf.

Co-funded by the European Union

tactic used 57% of the time when online abuse is inflicted. Text-to-image generative AI models make it easier to generate realistic-looking images of women in scenarios and situations that they were not in or did not consent to. [15]

Citizens' rights can also be at risk. Microtargeting advertising in combination with generative AI tools such as ChatGPT have raised concerns about the potential misuse of large language models in scaling microtargeting efforts for political purposes. [16] The Cambridge Analytica Scandal showcased the potential of AI technologies to influence voter behaviour in the US elections of 2016 as well as in the Brexit referendum. Data harvesting used to create psychological profiles and targeted messages aimed at directing voters toward specific options or to abstain clearly demonstrate AI's potential impact on **democratic participation**. At the same time, the use of AI in decision-making raises concerns over **transparency and fairness**. The use of "black box" AI systems (i.e. algorithms that produce results without showing to the end-user how these results were reached) undermines the principles of traceability and explainability of AI and hinders the contesting of administrative decisions, impacting on the rights to good administration, **access to justice and effective remedies.**

The use of AI in decision-making that concerns access to social welfare and healthcare services can reinforce negative stereotypes and deepen the social divide. Since 2013, Dutch tax authorities have been using an algorithm to detect fraud in childcare benefits applications by parents or caregivers. To perform the assessment, the authorities used information about nationality as a risk factor and assigned higher risk to applicants with dual or non-Dutch nationality. As a result, approximately 26.000 parents and caregivers were falsely accused of fraud. 76% of them represented ethnic minorities and most came from low-income families.[17] People accused of fraud received substantial fines and those that could not pay in time also lost access to their bank accounts, which were seized by the government. In some cases, high debts also led to the loss of employment and housing and over 1500 children were forcefully removed from their homes[18]. This scandal represented a pivotal moment for increased awareness among the Dutch society of how the use of algorithms can impact fundamental rights and people's lives, and led to the resignation of the Dutch cabinet in 2021.

[15] "Your opinion doesn't matter, anyway" *Exposing Technology-Facilitated Gender-Based Violence in an Era of Generative AI*, UNESCO (2023), accessible at https://unesdoc.unesco.org/ark:/48223/pf0000387483.
[16] *The persuasive effects of political microtargeting in the age of generative artificial intelligence*, Almog Simchon, Matthew Edwards, Stephan Lewandowsky, *PNAS Nexus*, Volume 3, Issue 2, February 2024, pgae035, https://doi.org/10.1093/pnasnexus/pgae035.
[17] For elaborate information on the algorithm and its impacts, see this report of Amnesty Netherlands: https://www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/
[18] https://www.dutchnews.nl/2022/05/childcare-benefit-scandal-more-children-were-removed-from-their-homes/

Co-funded by the European Union

Benefits

As with all new technologies, AI comes with risks as well as benefits. Hence, it is important to regulate it in a way that ensures it is used ethically and with the appropriate oversight. Human-centric approaches can utilise its potential to process vast datasets quickly and efficiently, recognise patterns, and make predictions to provide viable solutions to some of humanity's most pressing problems. The UN has proclaimed its intention to promote the use of AI in order to further the 2030 Agenda and Sustainable Development Goals (SDGs) in no uncertain terms and on multiple occasions. [19] The EU Strategy on AI commits to the same goal and undertakes to *"promote the use of AI, and technologies in general, to help solve global challenges, support the implementation of the Paris Climate agreement and achieve the United Nations Sustainable Development Goals"*. [20]

From agricultural applications that optimize food production and distribution in environmentally friendly ways to combat food scarcity and food poverty, to scientific models that offer ways out of the current climate crisis, to systems that streamline public administration processes to improve access to essential services such as health, education, and social welfare, AI has the potential to create new avenues for collective progress and well-being. The field of **medicine** is rapidly adopting AI technologies, aiming to capitalise on the opportunities it provides for research, early detection and diagnosis, telemedicine, and the streamlining of menial tasks. [21] Cost reduction and remote monitoring can improve access to high-quality healthcare for individuals in remote or underserved areas, thus reducing inequalities in this crucial area. However, healthcare systems should be mindful of risks related to inaccuracies, security risks for special categories of data, and overlooking social variables.

AI can also be used to enhance public participation and civic engagement by **democratising access to knowledge** and helping to address information overload. Fact-checking algorithms powered by Natural Language Processing (NLP) systems can identify and flag false claims, contributing to accurate and manageable information and facilitating public debate. Automated content moderation tools, when designed with transparency and accountability, can reduce hate speech and promote healthy

[19] See, for example, ibid. 1, 2.
[20] Communication from the Commission, Artificial Intelligence for Europe, Brussels, 25.4.2018, COM(2018) 237 final, accessible at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN.
[21] *Pros & Cons of Artificial Intelligence in Medicine*, Drexel University, College of Computing and Informatics, July 21, 2021, accessible at https://drexel.edu/cci/stories/artificial-intelligence-in-medicine-pros-and-cons/; *Balancing The Pros And Cons Of AI In Healthcare*, Jesse Corn, Forbes, Dec 1, 2023, accessible at https://www.forbes.com/councils/forbesbusinesscouncil/2023/12/01/balancing-the-pros-and-cons-of-ai-in-healthcare/; *Internet of Medical Things (IOMT): Applications, Benefits and Future Challenges in Healthcare Domain*, Gulraiz J. Joyia, Rao M. Liaqat, Aftab Farooq, and Saad Rehman National University of Sciences and Technology, Islamabad, Pakistan, Journal of Communications Vol. 12, No. 4, April 2017.

Co-funded by the European Union

democratic dialogue. [22]The risks of manipulated information, microtargeting, and deep fakes can, however, reduce the effectiveness of these tools.

AI in governance and public administration can help resolve persistent problems of efficiency, reduce backlogs, and redirect the energy of public servants into meaningful tasks. These functions have the potential to promote **good governance** and **good administration** principles. When applied in the area of **justice** they can be pivotal in ensuring the speedy yet diligent resolution of claims, a key requirement of **fair trials**. Public administrations using AI tools should be mindful of the risks associated with the lack of human oversight and automated decision-making and avoid the use of black box AI which violates the principles of transparency and contestability.

[22] See, for example: *Understanding Counterspeech for Online Harm Mitigation*, The Alan Turing Institute, accessible at https://www.turing.ac.uk/news/publications/understanding-counterspeech-online-harm-mitigation.

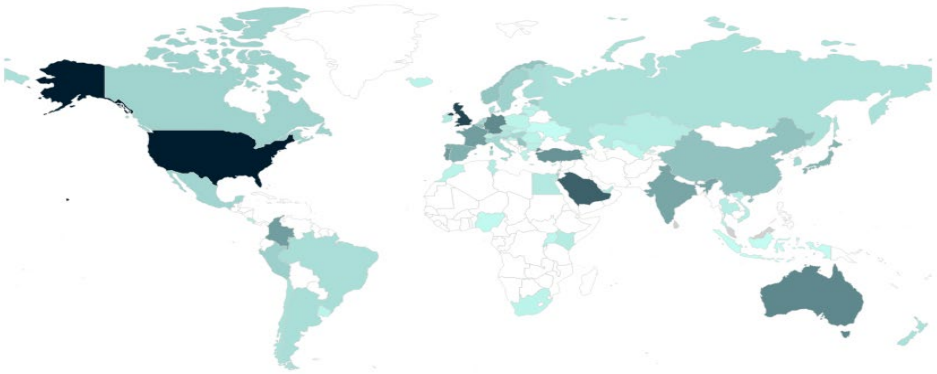
Co-funded by the European Union

International and EU approaches to safeguarding fundamental rights in the age of AI

General observations

As nations, regional bodies, and international actors increasingly recognize AI as a key driver of economic growth, technological advancement, and geopolitical influence, a global competition to lead in AI development has emerged. Countries are investing heavily in AI research, infrastructure, and deployment to secure competitive advantages in the global economy. At the same time, the evident ethical concerns over the deployment and use of these technologies have pushed legislators, policy-makers, and agencies with a relevant remit to adopt texts of varying legal force in an attempt to mitigate the risks and placate the voices against their widespread adoption. Since 2017, the Organisation for Economic Cooperation and development (OECD) has identified more than 1 000 policy and strategic initiatives [23] on AI from 69 countries and the EU.

Figure 2: Visualisation of AI initiatives by initiative count



Source: OECD

At the international level, the UN and the OECD play a critical role in shaping this balance. They provide platforms for dialogue, establish ethical frameworks, and set normative standards to guide the responsible development and deployment of AI. The UN emphasizes global cooperation and the integration of AI with the Sustainable Development Goals, advocating for technology that enhances equity and inclusivity, [24]while the OECD has developed principles promoting transparency, accountability, and human-centric, trustworthy AI, which have shaped AI-related policies worldwide. [25]

[23]OECD AI Policy Observatory, https://oecd.ai/en/dashboards/overview.
[24] Ibid. 1, 2.
[25] Recommendation of the Council on Artificial Intelligence, C/M(2019)10, adopted on 22/05/2019, accessible at http://web.archive.org/web/20240724092458/https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

Co-funded by the European Union

In Europe, both the EU and the CoE play a crucial role in the development and adoption of harmonized, legally binding regulatory initiatives, including the AI Act, and the Council of Europe (CoE) Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law. These instruments emphasise a risk-based approach to the regulation of AI.

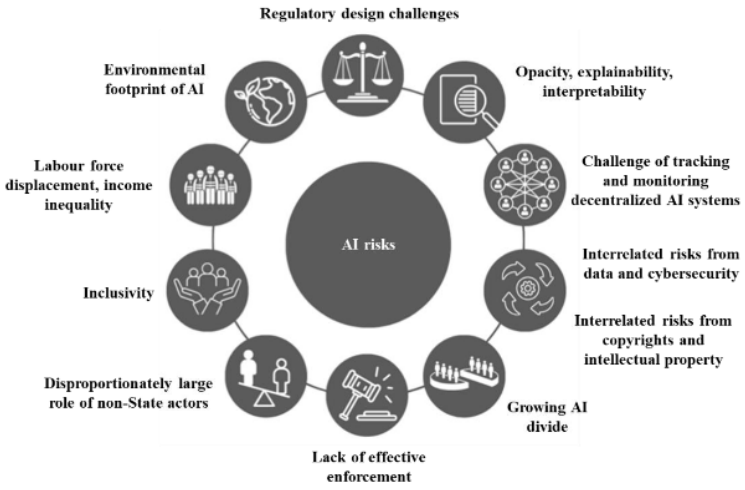
Unfortunately, despite the global efforts to mitigate AI-related risks, research shows that there is still a long way to go until an adequate level of fundamental rights protection is achieved, including in terms of appropriate enforcement mechanisms, equality and non-discrimination, workers' protection, and safety, security and reliability of AI systems. [26] This chapter examines different approaches to AI regulation and governance at the international, regional, and national levels, with a specific focus on fundamental rights protection. The different approaches are used to derive the common principles that guide policy initiatives and shape domestic and European laws.

International approaches

The United Nations

The United Nations (UN) has been addressing "frontier issues" such as artificial intelligence (AI) since 2017. [27] Recognizing AI's far-reaching impacts for peace and security, sustainable development, human rights, and humanitarian action, the UN emphasizes the importance of global AI governance, rooted in values-based frameworks like the UN Charter, international human rights law, and the 2030 Agenda for Sustainable Development.

Figure 3: Current and evolving global risks of artificial intelligence



Source: Inter-Agency Working Group on Artificial Intelligence (IAWG-AI)

[26] Adams, R., Adeleke, F., Florido, A., de Magalhães Santos, L. G., Grossman, M., Junck L. & Stone, K. (2024), *Global Index on Responsible AI 2024* (1st edition), South Africa, Global Centre on AI Governance.
[27] See, for instance, *2017 Chief Executives Board for Coordination (CEB) Survey on Frontier Issues*, accessible at https://unsceb.org/ceb-survey-frontier-issues.

Co-funded by the European Union

Resolution on Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development [28]

In March 2024, the UN General Assembly (UN GA) passed a Resolution on Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development. The Resolution relies on three key observations:

1. It recognizes that safe, secure and trustworthy AI systems have the potential to accelerate and enable progress towards the achievement of all 17 SDGs.
2. At the same time, it also recognises that AI may also inhibit this progress and create risks for human rights and fundamental freedoms, and reaffirms that these must be respected, protected and promoted online as they are offline, throughout the AI life cycle.
3. Finally, it draws attention to the need to minimise global inequalities in AI and include all nations, in particular developing ones, in regulatory action, as well as to enable them to overcome structural impediments and other obstacles they may face in fully accessing AI and harnessing its potential.

The resolution calls on states and other stakeholders to refrain from or cease the use of artificial intelligence systems that are impossible to operate in compliance with international human rights law or that pose undue risks to the enjoyment of human rights, especially of those who are in vulnerable situations. This risk-based approach underpins most policy and strategic initiatives launched in recent years.

The UNESCO Recommendation on the Ethics of Artificial Intelligence [29]

In line with its mission to bring people and nations together through education, culture and science, UNESCO produced, in November 2021, the **first-ever global standard on AI ethics** – the 'Recommendation on the Ethics of Artificial Intelligence', adopted by all 194 of its Member States. The Recommendation addresses (a) states, both as AI actors that use and deploy AI technologies, and as regulators of AI; and (b) public and private actors, in terms of ethical impact assessments.

The Recommendation is grounded on four core values: (a) respect, protection and promotion of human rights, fundamental freedoms and human dignity; (b) environment and ecosystem flourishing; (c) diversity and inclusiveness; (d) peaceful, just and interconnected societies; and establishes key ethical principles to uphold these values (see table 1).

To avoid becoming a theoretical exercise, the Recommendation also proposes concrete actionable policies for ethical AI, concerning the following areas: ethical impact assessment; ethical

[28] Ibid. 2.
[29] Ibid. 1.

Co-funded by the European Union

governance and stewardship; data policy; development and international cooperation; environment and ecosystems; gender; culture; education and research; communication and information; economy and labour; health and social wellbeing. Summaries of the key recommendations pertaining to fundamental rights follow below.

To assist with the implementation of its ethical framework, UNESCO has established:

- The Global AI Ethics and Governance Observatory, comprising knowledge and resources from around the globe, as well as information on states' readiness to adopt AI ethically and responsibly.
- An Ethical Impact Assessment tool, to assess compliance with AI ethics.
- A Readiness Assessment Methodology (RAM), which includes quantitative and qualitative questions designed to gather information about a country's AI ecosystem, including the legal and regulatory, social and cultural, economic, scientific and educational, and technological and infrastructural dimensions.

Table 1: Summary of Key Principles for the Ethical Use of AI in the UN framework [30]

Do no harm	● AI systems should not be used in ways that cause or exacerbate harm, including harm to social, cultural, economic, natural, and political environments. ● The Charter of the United Nations applies throughout the AI lifecycle, which should respect, protect and promote human rights and fundamental freedoms. ● The intended and unintended impact of AI systems should be monitored in order to avoid causing or contributing to harm.
Defined purpose, necessity and proportionality	The use of AI systems and specific AI method(s) employed for their operation should be justified, appropriate in the context and not exceed what is necessary and proportionate to achieve legitimate aims.
Safety and security	● Safety and security risks should be identified, addressed and mitigated throughout the AI system lifecycle. ● Robust frameworks should be in place to enable safe and secure AI systems, including a) sustainable, privacy-protected data access frameworks, b) appropriate safeguards against function creep, and c) fair and inclusive training, validation, and maintenance of AI models based on quality data.
Fairness and non-discrimination	● AI governance should promote fairness for equal and just distribution of the benefits, risks and costs, and to prevent bias, discrimination and stigmatization. ● AI systems should not lead to individuals being deceived or unjustifiably impaired in their human rights and fundamental freedoms.

[30] As derived from the *Principles for the ethical use of artificial intelligence in the United Nations system*, CEB/2022/2/Add.1, 27 October 2022.

Co-funded by the European Union

Sustainability	● Any use of AI should aim to promote environmental, economic and social sustainability. ● The impacts of AI should be assessed continuously and appropriate mitigation and/or prevention measures should be taken to address adverse impacts, including on future generations.
Right to privacy, data protection and data governance	● Privacy of individuals and their rights as data subjects must be respected, protected and promoted throughout the AI lifecycle. ● Adequate data protection frameworks and data governance mechanisms should be established or enhanced to ensure the integrity of the data used.
Human autonomy and oversight	● AI systems must not overrule freedom and autonomy of human beings. ● Human oversight and meaningful opportunities for human decision-making should be guaranteed throughout the AI lifecycle, including deciding when and how to use the AI system or to override its decisions.
Transparency and explainability	● Transparency and explainability should be ensured at all stages of AI lifecycle and in decision-making processes involving AI systems. ● Technical explainability requires that the decisions made by an AI system can be understood and traced by human beings. ● When a decision which may impact individual rights, fundamental freedoms, entitlements, services or benefits, is informed by or made based on AI algorithms, the person(s) concerned should be meaningfully informed in an understandable manner, including the reasoning.
Responsibility and accountability	● Appropriate oversight, impact assessment, audit and due diligence mechanisms, including whistle-blowers' protection, should be in place to ensure accountability. ● Appropriate governance structures should be established or enhanced which attribute the ethical and legal responsibility for AI-based decisions to humans or legal entities. ● Harms caused by and/or through AI should be investigated and appropriate action taken in response.
Inclusion and participation	● An inclusive, interdisciplinary and participatory approach, which promotes gender equality, must be taken when designing, deploying and using AI systems. ● Meaningful consultations should be conducted with all relevant stakeholders and affected communities.

Source: UN CEB

The future of AI within the UN framework

The UN 2.0 aims to combine the internal transformation of the UN with the transformation of its Member States to promote the SDGs through innovation and responsible AI. In September 2024, the Summit of the Future adopted the Pact for the Future, including a Global Digital Compact and

Co-funded by the European Union

Declaration on Future Generations. [31] The Compact sets specific objectives related to digital governance, including fostering an inclusive, open, safe and secure digital space that respects, protects and promotes human rights (Objective No 3), and enhancing international governance of artificial intelligence for the benefit of humanity (Objective No 5), to promote a balanced, inclusive and risk-based approach to AI governance.

Figure 4: The UN 2.0 Vision

Source: UN 2.0 Policy Brief No 11 [32]

The G7 and the OECD

In 2018, the G7 established the Global Partnership on AI (GPAI), a multistakeholder initiative bringing together leading experts from science, industry, civil society, international organizations and government with the aim to promote responsible AI in practice. GPAI has eventually grown to involve 29 international partners – including the EU – and three expert support centres. [33] In 2024, the GPAI and the OECD entered into an integrated partnership [34] to promote values and principles included into the OECD Recommendation on Artificial Intelligence. [35]

The OECD Recommendation is open to both members and non-members of the OECD, who have committed to adhere to the following principles:

- Inclusive growth, sustainable development and well-being
- Respect for the rule of law, human rights and democratic values, including fairness and privacy
- Transparency and explainability
- Robustness, security and safety, and

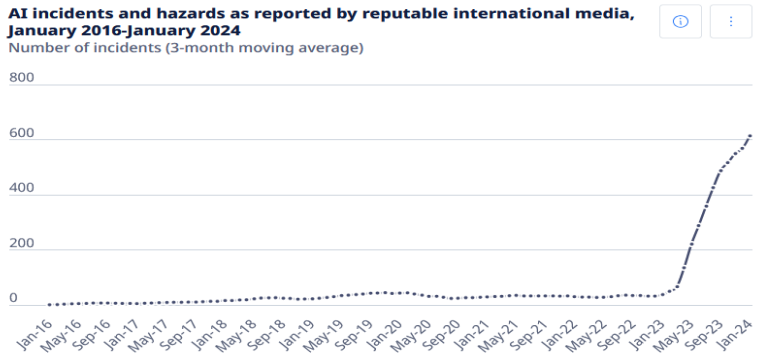
[31] Summit of the Future outcome documents (September 2024), *Pact for the Future, Global Digital Compact and Declaration on Future Generations*, available at https://www.un.org/en/summit-of-the-future/global-digital-compact (accessed on 25/11/2024).
[32] UN 2.0: Forward-thinking culture and cutting-edge skills for better United Nations system impact, Policy Brief 11 - Our Common Agenda, September 2023, accessible at https://www.un.org/two-zero/sites/default/files/2023-09/un-2.0_policy-brief_en.pdf.
[33] The International Centre of Expertise of Montreal for the Advancement of Artificial Intelligence (ICEMIA), the Paris Centre, piloted by INRIA, the NICT, in Tokyo, Japan.
[34] Announced during the 6th meeting of the GPAI Ministerial Council held on 3rd July 2024 in New Delhi.
[35] Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, accessible at http://web.archive.org/web/20240724092458/https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449.

Co-funded by the European Union

- Accountability

To support and promote the above principles, the OECD has established an Observatory of Policies, data and analysis for trustworthy artificial intelligence, an AI Incidents Monitor (AIM), and an updated definitions of relevant terms, used as a reference in various regional and national legal and policy documents, including the EU's Coordinated Plan on AI.

Figure 5: AI incidents and hazards as reported by reputable international media



Source: OECD, AIM

The Council of Europe Framework Convention on Artificial Intelligence

The CoE established in 2019 the Ad hoc Committee on Artificial Intelligence (CAHAI) with the aim to examine the feasibility of a legal framework for the development, design and application of AI. Upon fulfilling its mandate, the CAHAI was dissolved and succeeded by the Committee on Artificial Intelligence (CAI), which processed its proposals and drafted the first-ever international legally binding treaty, aimed at ensuring the respect of human rights, the rule of law and democracy legal standards in the use of AI systems, open for signature since 5 September 2024.

Despite being produced by a regional body, the Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, [36] has broader, global implications, as it is open for signature by non-member states that share the values and aims of the CoE, [37] and creates obligations in relation to AI systems used in the state – including companies acting on its behalf – as well as those used in the private sector, provided that states-parties opt for it via declaration, as recommended by the Parliamentary Assembly of the Council of Europe (PACE). The Convention emphasises the fundamental rights online-offline continuum and is guided by the principles of respect for human dignity and

[36] Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law, Council of Europe Treaty Series - No. 225, Vilnius, 5.IX.2024, accessible at https://www.coe.int/en/web/artificial-intelligence/the-framework-convention-on-artificial-intelligence.
[37] Non-European states that have signed the Convention so far include the United States of America and Israel.

Co-funded by the European Union

individual autonomy, transparency and oversight, accountability and responsibility, equality and non-discrimination, privacy and personal data protection, reliability, and safe innovation.

To ensure adherence to these principles, the Convention envisages concrete measures to be undertaken if not already in place:

a) Accessible and effective remedies for violations of human rights resulting from the activities within the lifecycle of AI systems;

b) Effective procedural guarantees, safeguards and rights where an AI system significantly impacts on the enjoyment of human rights, including notifying persons that they are interacting with AI instead of a human being;

c) Risk and impact management frameworks and mitigation measures to identify, assess, prevent, and mitigate risks to human rights, democracy and the rule of law, throughout the AI lifecycle.

Additional features include: the obligation to enact measures and mechanisms to identify content generated by AI systems; assess the need for a moratorium, a ban or other appropriate measures concerning uses of AI which may be incompatible with human rights standards; respect equality, including gender equality, the prohibition of discrimination, and privacy rights; ensure that AI systems are not used to undermine democratic institutions and processes. The Convention's provisions do not apply to actions related to the protection of national security interests or matters of national defence, but states are obliged to ensure that these respect international law, democratic institutions and processes. It also does not apply to research and development activities, except where they may have the potential to interfere with human rights, democracy or the rule of law.

Despite its significance as a legally binding instrument with a global reach, the Convention has also faced considerable criticism by the civil society, [38] the PACE, [39] and the UN, [40] for failing to reach specific standards in the following areas:

[38] See the Conference of International NGOs of the Council of Europe (CINGO) *Recommendations to PACE prior to drafting its Opinion in the Draft Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law*, accessible at https://ecnl.org/news/coe-and-ai-convention-last-chance-salvage-human-rights-blanket-exemptions-and-weak-language; the European Center for not-profit Law (ECNL)'s Reflections on the Council of Europe AI Convention, accessible at https://ecnl.org/news/council-europe-ai-convention-adopted-ecnls-reflections.

[39] PACE Draft Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law, Opinion 303 (2024), accessible at https://pace.coe.int/en/files/33517/html.

[40] UN High Commissioner for Human Rights on X (formerly known as Twitter) https://x.com/volker_turk/status/1790836806994264340; UN Special Rapporteur on Environmental Defenders under the Aarhus Convention, Statement on the proposed Council of Europe Framework Convention on Artificial Intelligence,
Human Rights, Democracy and the Rule of Law, 8 May 2023, accessible at https://unece.org/sites/default/files/2024-05/SR_EnvDefenders_Statement_CoE_FrameworkConvention_AI%20and%20Human%20Rights_08.05.2024.pdf.

Co-funded by the European Union

- The unequal application of its provisions to the public and private sectors, which, according to the PACE, contravenes the principle of the States' positive obligations to safeguard fundamental rights against private actors and creates a significant loophole.
- The fact that certain ethical principles should be formulated as positive individual rights rather than general principles (for instance, privacy, equality and non-discrimination).
- The blanket exception of AI use in cases involving national security.
- The failure to include the preservation of health and the environment as one of the general principles underpinning the activities within the lifecycle of AI systems.
- The omission of positive uses of AI for democratic processes, for instance improving government accountability and facilitating democratic action and participation.

In addition to the Framework Convention, and complementing its provisions, the CoE has adopted several sector-specific instruments, aiming to address issues with particular relevance to specific fields or AI uses, including the fields of Biomedicine and Health, Media and Information Society, the Rights of the Child, Social Cohesion, Crime, Democracy and Governance, Education, Justice, Personal Data, and others. [41] In addition, to facilitate the practical application of the Convention, and perhaps fill in some of the gaps identified above, the CAI has developed a Risk and Impact Assessment of AI Systems from the point of view of Human Rights, Democracy and Rule of Law (HUDERIA), a non-binding methodology, piloted during June-July 2024, and due to be launched at the end of the year.

The European Union

The European Union (EU) has been actively pursuing strategic and regulatory approaches to AI over the past years, grounded in respect for the fundamental EU values, established in Art. 2 of the Treaty on European Union (TEU) and the Charter, while also encouraging innovation and excellence. The foundation of EU policies in the area of AI are the 2018 EU Strategy for Artificial Intelligence for Europe [42] and the Coordinated Plan on AI, [43] published in 2018 and updated in 2021. In line with the relevant international approaches, the EU highlights the need to embrace change while also ensuring

[41] For a full list of CoE's relevant initiatives, see the *Overview of the Council of Europe activities in the field of artificial intelligence*, available at https://rm.coe.int/brochure-artificial-intelligence-en-march-2023-print/1680aab8e6 (accessed on 25/11/2024).
[42] Artificial Intelligence for Europe, Communication from the Commission, COM(2018) 237 final, Brussels, 25.4.2018, accessible at https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2018:237:FIN.
[43] Coordinated Plan on Artificial Intelligence, Annex to the Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions, Fostering a European approach to Artificial Intelligence Brussels, 21.4.2021 COM(2021) 205 final, accessible at https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review.

Co-funded by the European Union

an appropriate ethical and legal framework, in line with the EU Charter of Fundamental Rights, and an environment of trust and accountability around the development and use of AI. [44]

Source: European Commission

Figure 6: Key axes of the EU Coordinated Plan on AI

Pursuant to these initiatives, the EC has established and facilitated the work of the High-Level Expert Group on AI (AI HLEG), which produced the Ethics Guidelines for Trustworthy Artificial Intelligence, a checklist of principles and requirements of trustworthy AI, including technical and non-technical, and the Assessment List for Trustworthy AI (ALTAI), as a tool to support the actionability these requirements and guide developers and deployers of AI in implementing them in practice.

The AI HLEG Guidelines and the ALTAI emphasise seven principles, largely corresponding to those addressed in the previous sections: (1) human agency and oversight, (2) technical robustness and safety, (3) privacy and data governance, (4) transparency, (5) diversity, non-discrimination and fairness, (6) environmental and societal well-being and (7) accountability. They also draw attention to the fact that there might be fundamental tensions between the different principles and requirements identified and that there is a need to continuously identify, evaluate, document and communicate these trade-offs and their solutions.

The AI Act

The culmination of the above policies and strategic initiatives was the adoption in June 2024 of the AI Act. [45] The main objective of the new Regulation is to promote innovation and the uptake of human-centric and trustworthy AI, while ensuring a high level of protection of the health, safety, and

[44] The same principles are echoed in the 2023 European Declaration on Digital Rights and Principles for the Digital Decade, which contains specific provisions on the interaction of people and AI systems, including ensuring an adequate level of transparency and information about their use; the use of adequate datasets to avoid algorithm bias; and promoting trustworthy standards to ensure respect with fundamental rights.
[45] Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828, accessible at https://eur-lex.europa.eu/eli/reg/2024/1689/oj.

Co-funded by the European Union

fundamental rights, as well as democracy, the rule of law and the natural and cultural environmental, against the harmful effects of AI systems. In addition, the Regulation aims to promote a consistent and high level of protection throughout the Union and address divergences among the regulations adopted by individual MS, which may hamper the free circulation, innovation, deployment and the uptake of AI systems and related products and services within the internal market.

The AI Act has adopted a risk-based approach to the regulation of AI systems, whereby a "risk" is defined as "the probability of an occurrence of harm and the severity of that harm". It, therefore, prohibits placing in the single market, putting into service and using AI systems that are incompatible with human wellbeing, fundamental rights, and EU values, as listed in its Article 5. In addition, it contains rules for the classification of AI systems as high-risk (Article 6), as well as a list of specific high-risk systems (Annex III), for which providers (developers), deployers, importers and distributors, as well as third parties as appropriate, are subject to heightened obligations. In addition to AI systems posing unacceptable or high risks to fundamental rights, the AI act also regulates systems that pose specific transparency risks (e.g. chatbots, requiring that humans are aware that they are interacting with AI), as well as general-purpose AI models including large generative AI models that carry systemic risks due to their capacities or the large scale of their use.

The following AI systems are considered to contravene EU values because they violate fundamental rights and are, therefore, prohibited under Article 5 AI Act:

- Exploitation of vulnerabilities of persons, manipulation and use of subliminal techniques;
- Social scoring for public and private purposes;
- Individual predictive policing based solely on profiling people;
- Untargeted scraping of internet or CCTV for facial images to build-up or expand databases;
- Emotion recognition in the workplace and education institutions, unless for medical or safety reasons (i.e. monitoring the tiredness levels of a pilot);
- Biometric categorisation of natural persons to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs or sexual orientation. Labelling or filtering of datasets and categorising data in the field of law enforcement is still possible;
- Real-time remote biometric identification in publicly accessible spaces by law enforcement, subject to exceptions.

Despite the EC defining the exceptions to the use of prohibited AI practices as "narrow", [46] in reality, their application may render their prohibition null and void. For example, despite the prohibition

[46] EC Artificial Intelligence – Questions and Answers, accessible at https://ec.europa.eu/commission/presscorner/detail/en/qanda_21_1683.

Co-funded by the European Union

of real-time remote biometric identification in public spaces, the AI Act allows for their use for missing persons or victims of abductions, preventing terrorist attacks or identifying suspects of serious crimes. These exceptions leave a vast area of police discretion, as evidenced, among others, in past instances of abuse of counter-terrorism legislation. [47] Similarly, systems that purport to recognise emotions have only been banned in the areas of employment and education. Despite existing evidence of serious harm and doubtful scientific basis of 'emotion recognition', it will still be possible to use such systems in the area of law enforcement, migration or justice. [48]

AI regulators may also run on another key issue when it comes to establishing safeguards for fundamental rights, the lack of a coherent and comprehensive definition of what is included in the definition of an AI system. Article 3 of the AI Act defines it as "a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and 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". The broadness, and consequent vagueness of this definition, which can create loopholes for abuses, has led to the Commission adopting a set of detailed guidelines on its various components. [49]

High-risk AI systems, according to Article 6 AI Act, are (a) those that are safety components of products, or that are products themselves, which are covered by the Union harmonisation legislation referred to in Annex I (e.g., machinery, toys, equipment and protective systems, etc.) and are also required to undergo a third-party conformity assessment with a view to placing them on or putting them into service in the single market; and (b) the AI systems referred to in Annex III (biometrics, critical infrastructure, education and vocational training, access to and enjoyment of essential services and benefits, law enforcement, asylum, migration and border control, administration of justice and democratic processes), provided that they pose a significant risk of harm to the health, safety or fundamental rights of natural persons, including by materially influencing the outcome of decision making.

As it stands, the high-risk AI systems list of Annex III is exhaustive. However, the EC retains the power to amend it by either adding or modifying use-cases or by removing high risk systems, provided that the conditions of amendment contained in Article 7 AI Act are fulfilled.

[47] European Network Against Racism, *Suspicion, Discrimination and Surveillance: The impact of counter-terrorism law and policy on racialised groups at risk of racism in Europe*, Brussels 2021, accessible at https://www.enar-eu.org/wp-content/uploads/suspicion_discrimination_surveillance_report_2021.pdf.
[48] For more examples, see ECNL, *Packed with loopholes: why the AI Act fails to protect civic space and the rule of law*, 03-04-2024, accessible at https://ecnl.org/news/packed-loopholes-why-ai-act-fails-protect-civic-space-and-rule-law.
[49] https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application.

Co-funded by the European Union

The preamble to the AI Act (48) identifies certain Charter rights as the most at risk: human dignity, respect for private and family life, protection of personal data, freedom of expression and information, freedom of assembly and of association, non-discrimination, the right to education, consumer protection, workers' rights, rights of persons with disabilities, gender equality, intellectual property rights, the right to an effective remedy and to a fair trial, the right of defence and the presumption of innocence, and the right to good administration. In addition to those rights, the Act draws attention to the particular vulnerabilities of children as rights bearers, [50] as well as the right to environmental protection enshrined in Article 37 of the Charter, which is also related to the health and safety of human beings. Despite considering the area of asylum, migration and border control as high-risk, it fails to mention the rights to asylum and non-refoulement, enshrined in Articles 18 and 19 of the Charter.

The AI Act entered into force on August 1st 2024, while prohibitions kick in in February 2025. However, some requirements on the high-risk AI systems will only be applicable at the end of a transitional period, in August 2nd 2026, while AI used by the EU's migration-related databases does not need to be compliant with the Regulation until 2030. To mitigate some of the adverse effects which may take place in the interim, the EC is promoting the AI Pact, which aims to incentivise the industry to adopt the AI Act's requirements proactively, before the legal deadline. [51] The Pact is signed by more than 130 companies, including large corporations, Small and Medium Enterprises (SMEs) from the EU and beyond.

Despite the importance of regulating AI at the EU level, the resolve shown, for instance, in the adoption of the General Data Protection Regulation (GDPR) is missing here. The EU has opted for a risk-based instead of a rights-based approach, seemingly caving to industry interests and capitalising on the public sentiment that favours safety and security. General, abstract provisions and far-reaching exceptions further erode its protections and render it vulnerable to violations. Some key concerns include: [52]

- Gaps and loopholes in relation to prohibited uses of AI, mentioned above, risk rendering the relevant provisions purely declaratory. In addition, Member States and national authorities will likely be left with the task to close these loopholes, leading to a fragmented application of the AI Act, limiting its effectiveness to coherently regulate AI within the single market.

[50] In accordance with Article 24 of the Charter and the United Nations Convention on the Rights of the Child (UNCRC), including in the UNCRC General Comment No 25 as regards the digital environment.
[51] For mor information see https://digital-strategy.ec.europa.eu/en/policies/ai-pact.
[52] For more detail, see ECNL, *Packed with loopholes: why the AI Act fails to protect civic space and the rule of law, an analysis of the AI Act from the rule of law and civic space perspectives*, accessible at https://ecnl.org/news/packed-loopholes-why-ai-act-fails-protect-civic-space-and-rule-law.

Co-funded by the European Union

- Risk assessment is left to the AI companies or the authorities using the relevant systems, potentially creating conflicts of interest and reducing transparency and accountability. Crucially, paragraph 3 of Article 6 AI Act, leaves the assessment on whether an AI system does or does not pose significant risks of harm to the health, safety or fundamental rights of natural persons, or materially influences the outcome of decision making to the AI providers. The responsibility to investigate all self-exempted AI systems would fall on the newly established national and EU authorities, which might lack the financial and human resources to do it effectively.
- Impact assessment requirements fall short to safeguard their effectiveness. Specifically, although AI deployers of high-risk systems must list potential fundamental rights impacts, they are not obligated to assess whether those impacts are acceptable or to put in place measures to prevent them (only to foresee mitigation measures in case they materialise). In addition, there is no requirement to consult external stakeholders – such as the civil society or people affected by the AI system. Finally, law enforcement and migration authorities are under no obligation to publish summaries of the findings of impact assessments they conduct, in contrast with private actors, thus preventing public oversights and scrutiny, raising concerns related to transparency, the rule of law, and the protection of the civic space.
- Regulatory loopholes, such as national security and law enforcement exemptions, could be exploited to weaken democratic institutions and processes and the rule of law. The blanket exception of activities related to national security from the Act's scope of application creates opportunities for governments to invoke it to use otherwise prohibited systems, for example in the area of surveillance. [53]
- Civic participation in the implementation and enforcement of the AI Act is not sufficiently guaranteed. In addition to the opacity and lack of external evaluation of impact assessments, referred to above, CSOs may only represent individuals whose rights have been violated when their case concerns consumer rights. The only formal way for civil society to participate in the implementation and monitoring of the AI Act will be through membership in the advisory forum to the newly established AI Office and AI Board.

Other applicable legislation

Despite its specific scope, the AI Act is far from the only piece of EU legislation that applies to the development, deployment and use of AI systems. The EU framework on fundamental rights, consisting

[53] The "national security" excuse has been used repeatedly to justify covert surveillance of citizens and politicians by the National Intelligence Agency in Greece, instigated in part by the European Parliament's PEGA Committee as regards the use of illegal spyware.

Co-funded by the European Union

mainly of the Charter and the European Convention on Human Rights (ECHR), as well as other CoE and international human rights instruments are also applicable, including the CoE 108+ Convention for the protection of individuals with regard to the processing of personal data, the 1966 International Covenant on Civil and Political Rights; the 1966 International Covenant on Economic, Social and Cultural Rights; the 1965 International Convention on the Elimination of All Forms of Racial Discrimination; the 1979 Convention on the Elimination of All Forms of Discrimination against Women; the 1984 Convention against Torture; the 1989 Convention on the Rights of the Child; the 2006 Convention on the Rights of Persons with Disabilities; and the 2006 International Convention for the Protection of All Persons from Enforced Disappearance. [54]

In addition to international and regional human rights instruments, secondary EU legislation, including the GDPR, the anti-discrimination legislation, [55] the directives on unfair commercial practices, [56] the Digital Services and the Digital Markets Acts, [57] and other EU laws containing provisions relevant to the protection of fundamental rights, continue to apply and may lead to the prohibition of AI practices, even if these are not covered by the AI Act. Additional protections adopted at the domestic level may also apply in cases outside the scope of the Act (e.g., for algorithmic systems not strictly fulfilling the definition of AI contained in the Act).

The EC AI governance framework

Before moving on to examine concrete use cases of AI systems that pose risks or, indeed, create opportunities for fundamental rights, it is useful to take a closer look on how the European Commission, the EU's largest administration, regulates the use of AI in its own offices and the types of AI systems it uses or plans to use in the near future.

The EU AI system of governance is established in Chapter VII of the AI Act (Articles 64-69It includes the European AI Office, operating within the EC with a mandate to develop Union expertise and capabilities in the field of AI, support the AI governance bodies established in the EU Member States and enforce the rules for general-purpose AI models, including requesting information and measures from providers, and applying sanctions.

[54] For more on the applicable international norms, see De Schutter, O. (2015), *International Human Rights Law: Cases, Materials, Commentary*, Cambridge University Press, 2nd edition.
[55] Including Directives 2000/43/EC on equal treatment irrespective of racial or ethnic origin; 2000/78/EC on equal treatment in employment and occupation; 2004/113/EC on the equal treatment of men and women in the access to and supply of goods and services; and 2006/54/EC on the equal treatment of men and women in employment and occupation (recast).
[56] Including Directives 2005/29/EC on unfair commercial practices; (EU) 2019/2161 on the better enforcement and modernisation of Union consumer protection rules; and (EU) 2024/825 on empowering consumers for the green transition through better protection against unfair practices and through better information, applicable from 2027.
[57] Regulations (EU) 2022/2065 and (EU) 2022/1925.

Co-funded by the European Union

The AI Office is advised and assisted by the European Artificial Intelligence Board, comprising one representative with the requisite expertise on AI matters per Member State, acting as a contact point and facilitating coordination among national authorities, as well as between them and the AI Office. The AI Board is also tasked with facilitating harmonisation of administrative practice and may issue recommendations and opinions on matters relevant to the implementation of the AI Act. The Board will further be assisted by an advisory forum providing technical expertise, comprising permanent and non-permanent members. [58] The permanent members are the Fundamental Rights Agency of the EU (FRA), the European Union Agency for Cybersecurity (ENISA), the European Committee for Standardization (CEN), the European Committee for Electrotechnical Standardization (CENELEC), and the European Telecommunications Standards Institute (ETSI). Non-permanent members are selected - in a balanced way - among stakeholders representing commercial and non-commercial interests, e.g., the industry, start-ups, SMEs, civil society and academia. Finally, the EC is currently in the process of drafting the implementing act for the establishment of the Scientific Panel of Independent Experts, provided for in Article 68 AI Act, to support the implementation and enforcement of the AI Act, and provide advice and support to the AI Board, market surveillance authorities, the AI Office, and individual Member States.

The Commission has issued a Communication [59] presenting its strategic vision on the use of AI and detailing the areas and purposes where it uses or plans to use it, as well as the potential risks and mitigating measures it has undertaken to ensure the obligations contained in the AI Act are fully upheld in its own activities.

The Commission has identified four priority areas where implementing AI technologies will bring the highest benefit, namely (i) enhancing document summarization capabilities, (ii) streamlining the preparation of briefings and responses to questions, (iii) introducing a conversational platform that supports non-classified human-like dialogues, and (iv) providing generative AI services to leverage the vast data, information and knowledge base that the administration has across various business areas.

AI systems already in service include:

- **eTranslation** and **eSummary**: an AI-powered language services that provide automated translation and summaries, both to the EU institutions, bodies, and agencies and to other users in the 24 official languages of the Union and other geopolitically or socio-economically relevant languages.

[58] Yet to be set up by the time of writing.
[59] *Artificial Intelligence in the European Commission (AI@EC), A strategic vision to foster the development and use of lawful, safe and trustworthy Artificial Intelligence systems in the European Commission*, Communication to the Commission, Brussels, 24.1.2024, C(2024) 380 final, available at https://commission.europa.eu/system/files/2024-01/EN%20Artificial%20Intelligence%20in%20the%20European%20Commission.PDF.

Co-funded by the European Union

- **Publio**: an AI-powered service for supporting users in their discovery of EU law and EU publications, thus also contributing to greater accessibility.
- **Doris (Data Oriented Services) drive-in**: a system that provides sentiment analysis, keyword extraction, summarisation, and named-entity recognition to semi-automatically analysis in any type and document. There is also specific dashboard for public consultations answers (Doris public consultation dashboard).
- **SeTA (Semantic Text Analyser):** applications built on SeTA are successfully in use for metadata creation, document classification and discovery. SeTA is being tested for other use cases.
- **AI systems currently in development and testing include:**
- **eBriefing**: AI-powered language service that produces topic-based overviews or briefings from a given set of relevant input documents.

- **EC Conference – speech-to-text**: speech-to-text technologies to enable multilingual captioning of speakers and subtitling of live web streaming or audio-visual material to improve access to content for people with disabilities.
- **Competition case management**: AI facilitating efficient search and analysis of documents and other data, enhancing the investigative process. The EC deems this as particularly important for ex officio investigations and soon for the enforcing the Digital Markets Act (DMA).
- **Fraud detection/cybersecurity/disinformation**: the Commission is working on incorporating AI systems that will better assess risks and target controls, including border controls; detect fraud and illicit activities related to the safety of products, services (especially those sold online) and tax compliance.
- **Science for policy**: AI is used to make available scientific evidence for EU policy making, including digital policies and other areas, such as agriculture, crisis management, security, transport, health, and consumer protection.
- **Fostering internal digital transformation**: generative AI used to test leading generative AI models in a safe environment and explore their potential uses to build new applications and services.
- **Improving the user experience on the Funding and Tenders Portal**: new AI-powered tools will be deployed to launch an advanced search engine based on NLP, enabling users to find new funding opportunities by using concepts rather than matching keywords. There will also be a recommendation system aimed at providing users with notifications in their area of expertise.

Co-funded by
the European Union

- **Complaints Handling with AI (CHAI)**: generative AI LLM technologies aim to help case handlers deal with complaints in shorter timespans, through 'smart search' and 'smart drafting' capabilities allowing for re-use of past replies from similar complaints.
- **Specialised IT systems:** AI systems like the ATHINA IT will be used to support organised and systematic collection, analysis, and interpretation of information/data from all sources using AI tools to detect, verify and investigate potential cross-border health risks/emergencies and to rapidly respond by ensuring the development, manufacturing, procurement, and equitable distribution of key medical countermeasures.

In addition to the systems currently use and those whose deployment is imminent, the AI@EC Communication also includes areas where the deployment of AI systems is under consideration. These include: systems to support the legislative process, policy monitoring and responses to parliamentary questions; the drafting of non-sensitive content for briefings, reports, and other documents; people-focused HR services; financial processes; project-proposal assessment; detection of risk, fraud, and unethical situations; threat intelligence; speech-to-text conversion; engagement with the public and stakeholders; a greener Commission in combination with internet of things; European statistics; exploring opportunities to improve other search functionalities within the Commission; and corporate-level AI services, such as conversational platforms and generative AI tools, to integrate knowledge bases across policy areas.

The AI@EC Communication preceded the finalisation and adoption of the AI act and anticipated its content. The EC adopted the risk-based approach to AI, fostered by the AI Act, and put in place measures and procedures complying with and operationalising the forthcoming rules of the AI Act. Thus, the Communication foresaw measures regarding the introduction of operational and pragmatic guidelines for the Commission staff (which are yet to be set up at the time of writing); the classification and assessment of AI systems already in use or whose use is planned; the commitment to refrain from using AI systems that are incompatible with European values or that represent a threat to the security, safety, health, and fundamental rights of people; and a governance framework that will allow the Commission to fulfil its obligations in relation to AI (now included in the AI Act).

In light of the Ai@EC Communication, and prompted by the forthcoming and transparent approach of the EC, the European Ombudsman has initiated a Strategic Initiative, aiming to further

Co-funded by
the European Union

clarify the information provided. The initiative concluded with a closing note by the Ombudsman, issued in December 2024. [60]

The Ombudsman requested information specifically on three use cases she considers to have a direct impact on the right to good administration – whose safeguarding constitutes the main mandate of her Office – as well as their specific functionalities and general risks entailed for this right: a) the analysis of feedback from the public; b) the management of competition cases; and c) the handling of infringement complaints. The Commission noted in its responses to the Ombudsman that these three AI systems are still in the development or procurement phase and that it will decide on their use only after the implementation of all technical features and a period of testing by selected staff are completed.

As the adoption of the AI Act took place while the initiative was still ongoing, the Ombudsman also sought information on how the new rules will affect the Commission's own development and use of AI.

The Ombudsman's concerns regarding the use of the above AI systems may broadly be categorised in three areas, often intersecting with one another:

1. **External consultation, transparency and public participation.**

   The Ombudsman applauded the EC's commitment to transparency, showcased by its proactive communication on its use of AI, and urged it to provide periodic updates to maintain a high level of transparency in this area. However, she drew attention to the fact that the Commission publishes relevant information based on the "maturity status" of the systems, and admits that although public consultation could provide valuable insights, it is not a requirement when it comes to using AI to enhance its operations.

   Further, the Ombudsman noted that, given the rapid development of AI technologies, human oversight may not be a sufficient safeguard in the near future, as it cannot be reasonably expected of humans to repeat the reasoning process carried out by AI and to ensure its correctness, which would also defeat the goal of efficiency. It, therefore, may be necessary to carry out some form of public consultation before projects reach full maturity, as the public should have a say on whether the EU administration can and should introduce powerful AI tools that will potentially influence decision-making processes.

2. **Application of the AI Act's risk-based approach and good administration.**

   The Commission has highlighted in the AI@EC Communication that it is an early adopter of the risk-based approach fostered in the AI Act. In their meetings with the Ombudsman, EC

[60] European Ombudsman, Closing Note on the Strategic Initiative on how the European Commission decides on and uses artificial intelligence (SI/4/2024/MIK), 6/12/2024, accessible at https://www.ombudsman.europa.eu/en/doc/correspondence/en/196934.

Co-funded by the European Union

representatives confirmed that AI systems used by the Commission, which are not deemed 'high-risk' will not necessarily have to comply with the AI Act's requirements for high-risk systems. In this respect, the Ombudsman noted that the AI Act's requirements on high-risk AI systems "*largely correspond and give effect to general principles of good administration*" that the EU administration must uphold when using any AI systems, not only those falling under the 'high-risk' category of the AI Act, which only applies to some of the AI systems used. The right to good administration should be specifically considered in relevant impact assessments. Additional concerns in this area include that many of the AI Act requirements concern the development of the systems only, rather than their use; that the requirements will in principle apply only from 2030; and that specific systems used in decision-making processes may not fall under the definition of AI contained in the AI Act.

Table 2: Requirements for high-risk systems (Art. 9-15 AI Act) and corresponding principles of good administration

Requirements under the AI Act (for high-risk systems)	**Principles of good administration**
- maintain risk management solutions for high-risk AI systems - draw up and keep up to date technical documentation concerning these systems - enable relevant events to be recorded while the systems are operating (through 'logs') - ensure cybersecurity	- exercise diligence, care and caution
- train the high-risk AI systems on relevant, representative, error-free, and bias-free data - ensure that the outputs of the AI systems are accurate and robust	- take decisions based on complete, accurate, reliable, and consistent evidence
- ensure that the outputs of the AI systems can be interpreted by humans	- give reasons for any decisions
- ensure that the outputs of the AI systems are subject to human oversight	- assume full responsibility for decisions (no unilateral and unlimited delegation of decision making to third parties)

Source: European Ombudsman

3. **Use of AI in decision-making.**

Finally, the Ombudsman raised concerns regarding the use of AI systems for competition case and complaints management. Despite the Commissions assurances that its AI systems are meant to support but not replace human decision makers, and that case handlers will remain in full control of the draft replies to complainants, the Ombudsman stressed that depending on their functionalities, the AI tools used in these cases may in fact '*semi-automate*' decision making in

Co-funded by the European Union

these areas by influencing the choices made by human case handlers or decision-makers – a possibility recognized also in the AI Act.

Using AI in the context of decision making can also be problematic in terms of the competences awarded to the EC under EU law regarding the delegation of power to other entities – in this case, powerful AI systems, capable of materially influencing decision-making processes. [61]

In light of the previous analysis, it is clear that many of the concerns shared by the Ombudsman correspond to the criticisms raised in relation to the application of the AI Act in the Member States. The Commission diverges from past approaches implementing stricter requirements to EU bodies, to serve as an example of good administration and foster trust. [62] In fact, it seems determined to adopt all allowances made by the AI Act, limit public participation to the minimum, and only concern itself with high-risk systems, ignoring other categories, even among those included in the AI Act, such as powerful systems capable of carrying systemic risks. This is especially concerning given the vast scope of influence EC decisions may have for the entirety of the Union. Nevertheless, as reminded by the Ombudsman, the Commission is bound by other fundamental rights-related legislation and by EU primary law (the Treaties and the Charter) in the same way Member States are – if not held to a higher standard. The principles of good administration will ultimately define acceptable uses of AI, if the above concerns are not addressed internally, e.g., through a robust code of conduct, as suggested by the Ombudsman.

Potential use of AI by other EU bodies

While the EC has been forthcoming with information on its use of AI, both current and planned, as well as speculated, the same cannot be said about other EU bodies. In some high-risk areas, such as asylum, migration, border control, and policing, a notable lack of transparency in relation to fundamental rights obligations, [63] coupled with a clear investment in research related to the potential

[61] See Court of Justice, Case C-270/12, *UK v Parliament and the Council*, https://curia.europa.eu/juris/document/document.jsf?text=&docid=146621&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=7732888, paragraphs 41-42 and 53, according to which discretionary powers with a wide margin of discretion cannot be delegated, while technical powers must be precisely delineated in legislative provisions and be subject to an effective oversight.

[62] See for instance, Regulation 2018/1725 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies.

[63] See, e.g., Decision in OI/4/2021/MHZ on how the European Border and Coast Guard Agency (Frontex) complies with its fundamental rights obligations and ensures accountability in relation to its enhanced responsibilities; Decision on how the European Border and Coast Guard Agency (Frontex) assessed the human rights impact before providing assistance to non-EU countries for developing surveillance capabilities (case 1473/2022/MHZ); Decision on the refusal by the European Border and Coast Guard Agency (Frontex) to give public access to documents concerning a search and rescue operation (case 1610/2021/MIG); Decision on how the European Border and Coast Guard Agency (Frontex) complies with its fundamental rights obligations with

Co-funded by the European Union

of AI technologies, [64] has created concerns over the potential use of AI without the necessary guarantees in place to ensure protection of fundamental rights and compliance with ethical norms.

In the area of border control and management, **Frontex** has commissioned a study on Artificial Intelligence-Based Capabilities, [65] which examines nine different types of technologies looking at both their current and future desired state, as well as the requirements and risks to their adoption. These technology areas are: automated border control, maritime domain awareness, machine learning, optimization, surveillance towers, heterogeneous robotic systems, small autonomous unmanned aerial systems (sUAS), predictive asset maintenance, object recognition, and geospatial data analytics. Some of these technologies are already in use, while others are still under development. While the study emphasises the positive aspects of these technologies and how they can boost the effectiveness of border control operations, it also acknowledges important barriers and risks, including unfamiliarity with technology and uncertainty concerning its performance, financial cost, additional infrastructure requirements, data protection and regulatory barriers, limits on access to relevant technologies, and, finally, *"insufficient political or public acceptance (e.g. due to ethical and human rights concerns)"*.

Similarly to Frontex **Europol** has recently issued a Report on AI and Policing, [66] detailing AI's potential for law enforcement. In the words of Europol's Executive Director, Catherine De Bolle, "*Artificial intelligence will profoundly reshape the law enforcement landscape, offering unprecedented tools to enhance our ability to safeguard public safety. Europol is committed to staying at the forefront of these technological advancements*". [67]

The report focuses on the opportunities of AI to enhance law enforcement capabilities; improve operational efficiency; provide real-time insights in crisis situations; and improve international cooperation. It also draws attention to certain key requirements for the operation of AI systems, including ensuring appropriate technical infrastructure and expertise, successfully navigating legal and ethical challenges, and investing in bias mitigation and learning. It further discusses alignment with the

regard to search and rescue in the context of its maritime surveillance activities, in particular the Adriana shipwreck (OI/3/2023/MHZ).

[64] See, e.g., Artificial Intelligence-Based Capabilities for the European Border and Coast Guard – Final Report, Warsaw 2021, accessible at https://www.frontex.europa.eu/assets/Publications/Research/Frontex_AI_Research_Study_2020_final_report.pdf; Europol (2023), AI and policing - The benefits and challenges of artificial intelligence for law enforcement Europol Innovation Lab observatory report, Publications Office of the European Union, Luxembourg, accessible at https://www.europol.europa.eu/publication-events/main-reports/ai-and-policing; https://www.europol.europa.eu/media-press/newsroom/news/how-ai-can-strengthen-law-enforcement-insights-europols-new-report.

[65] Ibid. 62.

[66] Ibid. 62.

[67] https://www.europol.europa.eu/media-press/newsroom/news/how-ai-can-strengthen-law-enforcement-insights-europols-new-report.

Co-funded by the European Union

requirements of the AI Act, including the 'black box' issue and real-time biometric identification, as well the obligations related to impact assessments, authorization, and national laws.

To promote the use of artificial intelligence in a transparent and accountable manner, Europol participates in the project Accountability Principles for Artificial Intelligence (AP4AI), jointly implemented by the Centre of Excellence in Terrorism, Resilience, Intelligence and Organised Crime Research and the Europol Innovation Lab. [68]

[68] https://www.europol.europa.eu/media-press/newsroom/news/new-accountability-framework-to-use-artificial-intelligence-in-transparent-and-accountable-manner.

Co-funded by the European Union

National approaches

The AI Act was enacted partly due to the emerging initiatives by many Member States in the field of AI regulation and governance, with a view to promote harmonization of practices across the single market. However, the loopholes mentioned earlier in this document leave considerable leeway for domestic regulation in fields beyond its scope (e.g., in some of the excluded areas, such as national security, or for systems which may not strictly fall within the definition of AI adopted in the EU Regulation). Furthermore, a closer look at the current state of affairs prior to its provisions becoming directly applicable can shed light on the progress made and the steps required to achieve compliance in the five consortium Member States.

Bulgaria

Bulgaria's policies that promote the development and deployment of AI-based systems are anchored in its overarching national strategy, "Digital Transformation of Bulgaria, 2020–2030". This national strategy recognizes digitalisation as an important driver of economic growth, innovation, and modernization. It focuses on building digital infrastructure, fostering public-private partnerships, and enhancing the technological capacity of the workforce. The development of AI is expected to benefit multiple areas of action covered in the strategy such as education, digital economy, e-government. . The Strategy is complemented by the "Concept for the Development of Artificial Intelligence until 2030", which emphasises facilitating the uptake and use of AI, through research and innovation, human and technical capacity building, improved access to and collection of data, and investments, including to support the activities of Small and Medium Enterprises (SMEs). Added to these objectives is the aim to establish an appropriate regulatory framework for AI in line with the existing international legal and ethical standards and good practices.

Thus far, the latter objective has mostly been served through soft law and non-binding initiatives. For example, a Working Group on Artificial Intelligence was set up in early 2024, bringing together experts from government and civil society, industry, and academia, promoting public-private partnerships and civic participation in line with Measure No 4 of the Fourth National Action Plan under the Open Government Partnership Initiative. The Group's mandate includes the promotion of effective public dialogue on the development of standards for the use of AI to guarantee equal access and compliance with human rights. Its work "may" inform legislative initiatives. The Working Group plays a key role in public policy-making on the governance of AI. It can also contribute to legal initiatives which then need to pass through parliament before becoming law.

Co-funded by the European Union

Another soft-law initiative concerns the adoption of guidelines on the use of AI systems in education, [69] aiming to harness the potential of AI in an educational context, while mitigating risks to fundamental rights, as well as countering potential misuse that may cause harm to children (e.g., manipulation, bullying, harassment). These guidelines were largely prompted by the growing popularity of generative AI tools and their publication coincided with the launch of BgGPT, the first open Bulgarian large language model for Bulgarian. The scope of these guidelines is limited to generative AI in formal schooling and as such do not cover the entire gamut of risks that AI technologies can pose to children.

With privacy and data protection accounting for a large portion of AI-related concerns, it is not surprising that data protection authorities are already playing a significant role in shaping the AI legal and policy framework across the EU. The Commission for Personal Data Protection (CPDP) has prepared and published awareness-raising electronic resources in the form of two brochures, [70] addressing Big Data analytics and the use of AI as they relate to personal data protection, as well as a portfolio of awareness-raising activities aimed at safeguarding children in the digital space. [71] The brochures discuss the risks and benefits of using Big Data and facial recognition, respectively, specifically addressing issues related to privacy, data protection, and social media surveillance and profiling, and outlining the relevant responsibilities of public authorities. Contemporary Threats and Challenges to Personal Data Protection in the light of the Trends in the Development of Artificial Intelligence and Facial Recognition Technologies is an online awareness-raising brochure which aims to shed light on the benefits and risks of using facial recognition.[72] The brochure briefly discusses AI and what AI-enabled facial recognition is, noting that this technology is widely used for authentication purposes. AI-enabled facial recognition offers important benefits in the area of security of public spaces. However, the brochure also notes the risks associated with the use of this technology and, in particular the risk of privacy breaches and algorithmic bias and discrimination. The brochure makes a number of

[69] Bulgaria, "The Ministry of Education and Science Releases Awareness-Raising Guidelines for Teachers on the Use of AI in Education", *Press release*, Ministry of Education and Science, 16 February 2024. See also Bulgaria. Ministry of Education and Science (2024) "Guidelines on the Use of Artificial Intelligence in the System of Education".

[70] Bulgaria, Commission for Personal Data Protection (2022) "Big Data and Their Potential for Profiling – Brochure for Natural Persons". Commission for Personal Data Protection (2022) "Big Data and Their Potential for Profiling – Brochure for Personal Data Administrators"; "Contemporary Threats and Challenges to Personal Data Protection in the light of the Trends in the Development of Artificial Intelligence and Facial Recognition Technologies".

[71] See Bulgaria, Commission for Personal Data Protection, "Fundamentals of Data Protection Online". Bulgaria, Commission for Personal Data Protection, "Privacy in the Digital Age". Bulgaria, Commission for Personal Data Protection, "Rights of Children and Young People when Using Digital Platforms".

[72] Bulgaria, Commission for Personal Data Protection (2022) "Contemporary Threats and Challenges to Personal Data Protection in the light of the Trends in the Development of Artificial Intelligence and Facial Recognition Technologies".

Co-funded by the European Union

recommendations with regard to the responsible use of AI-enabled facial recognition. Key points include ensuring that personal data used in the development of such technologies must be extracted and processed in line with the GDPR requirements; that algorithms used for facial recognition comply with the existing rules for privacy, data protection, and non-discrimination; that humans must exercise oversight on AI systems; and that data manipulation is effectively countered, in order to ensure that facial recognition is not misused for the purposes of biased profiling.

Bulgaria has invested in initiatives to promote digital literacy and raise awareness about AI's implications to empower citizens to understand and challenge AI systems. The GATE Institute, the first dedicated Big Data and AI Centre of Excellence in Eastern Europe, [73] develops research capacity and potential in Big Data and AI and provides advanced infrastructure in terms of platforms, data, services and testing and experimentation facilities. A priority research area at GATE is the study of online disinformation in the Balkans to tackle social susceptibility to conspiracies and misinformation by developing new AI methods for large-scale disinformation analysis in English, Bulgarian and other Balkan languages, and acting as an amplifier of national and regional actions against disinformation. [74]

These efforts are complemented by capacity-building programs for public officials responsible for the oversight of implementing and regulating AI systems by the Institute of Public Administration (IPA) and the National Institute of Justice (NIJ), responsible for training Bulgarian civil servants and law enforcement and justice professionals, respectively. The IPA offers interdisciplinary courses on the use of AI systems, [75] NIJ has hosted webinars focusing on the application of AI tools in the investigation and prosecution of crimes and administers a discussion dashboard on the application of AI systems and tools in the legal practice. The dashboard is hosted on the NIJ's online platform dedicated to judicial ethics. [76]

The use of AI tools in Bulgaria is in an upward trajectory. BgGPT is a non-commercial LLM, trained on the specifics of Bulgarian language. It was developed by the Institute for Computer Science, Artificial Intelligence, and Technology (INSAIT) in Bulgaria, [77] and is used widely, across sectors. Legal professionals can also use local machine-learning based tools for case law analytics. The ADELE tool provides free access to judicial decisions issued by Bulgarian courts in two legal areas, namely, Value Added Tax (VAT) – exemptions and deductions, and Trademarks and Patents – claims for infringement.

[73] Co-funded by the EC and the Bulgarian Government and established as a joint initiative with Chalmers University of Technology and Chalmers Industrial Technologies, Sweden
[74] See "Annual AI Report Bulgaria 2022", AI Bulgaria and See News; also see, GATE: Intelligent Government Programme.
[75] Bulgaria, Institute of Public Administration (2024), Catalogue of Training Courses.
[76] Bulgaria, National Institute of Justice (2024) Judicial ethics platform: Information technologies in law.
[77] INSAIT, "INSAIT announced BgGPT – the first open AI model for Bulgarian language", *Press release*, 16 January 2024. See also BgGPT.

Co-funded by the European Union

The tool blends machine learning and natural language processing to extract knowledge and predict the outcomes of legal cases. [78] The Lexebra tool is an AI-based tool that focuses on case law of the Supreme Court of Cassation in Bulgaria emphasising decisions that aim to promote standardisation. Lexebra features over 250,000 selected cases from civil, commercial, and criminal case law, as well as case law of the Supreme Administrative Court. [79] Despite its many benefits in terms of speed and efficiency, automated case law analytics also has limitations. For example, depending on the algorithm structure relevant case law might not show up immediately, or might not feature as a top result, leading to a narrower interpretation of the law.

Estonia

The Estonian state is an early adopter of digitalization, investing in e-government since the early 2000s. The e-State, [80] comprising digital identification and X-Road data exchange layer, an open-source, secure, and decentralised data exchange platform, have solidified public trust in new technologies, when deployed and controlled by the state. [81] Against this background, it is not surprising that Estonia is one of the first countries to consider AI governance and regulation.

An AI Taskforce, established in 2018, has brought together public and private sector representatives on the initiative of the Ministry of Economic Affairs and Communications and the Government Office with the mission to develop specific proposals on the areas that would benefit most from AI and the measures required to support its adoption. At the same time, it also conducted legal analysis to identify the potential needs for reform to ensure legal clarity and safety in the use of the newly adopted technologies. [82] The Estonian government adopted the Taskforce's techno-optimistic view of AI, [83] and launched its first national AI strategy for 2019–2021, followed by subsequent strategies for 2022–2023 and 2024–2026 [84] with the aim to position Estonia as a leading AI adopter globally. The strategies

[78] See ADELE Pilot Tool.
[79] See BGToll: Toll Service; Ministry of Interior, Fixed automated technical means for speed control.
[80] P. K. Tupay, M. Mikiver, "The Estonian e-state and challenges of regulating public sector digitalization". *Public Digitalisation in a legal perspective: Status, challenges and opportunities for Nordic-Baltic cooperation.* Nordic Council of Ministers, 13 May 2024. Available at: https://pub.norden.org/temanord2024-503/estonia.html
[81] Kantar Emor, Inimeste privaatsusõigused ja isikuandmete kaitsmine 2020, *Ministry of Justice*. Available at: https://www.just.ee/uuringud
[82] Information System Authority (*Riigi Infosüsteemi Amet*), Bürokratt: Ülevaade krattide süsteemist. Available at: https://www.ria.ee/riigi-infosusteem/personaalriik/burokratt
[83] It is indicative that AI solutions are referred to as "*kratts*" – the term "*kratt*" comes from Estonian folklore, where a "*kratt*" is a mythical creature or servant that accomplishes tasks for its master.
[84] Kratid, Visioon ja kavad. Available at: https://www.kratid.ee/kratt-visioon.

Co-funded by the European Union

implement the directions and goals prescribed by other national development plans, including the Digital Agenda 2030 [85] and Data and Artificial Intelligence White Paper 2024-2030. [86]

The Estonian policy framework on AI is clearly geared toward promoting and facilitating its uptake by the public sector, as a driver of innovation and growth. While the state does exhibit interest to ensure that this process is done safely, with due regard to fundamental rights, and in a way that preserves public trust on the e-State, it also clearly stresses the need to ensure that AI regulation "does not become a barrier to the development and deployment of AI". [87] The impact of this approach is showcased in the impressive uptake of AI by the public sector: while only four AI projects had been carried out by an equal number of public sector bodies at the time the first AI Strategy was being prepared, this number had increased to 80 by 2022, and by 2024, more than 130 AI projects had been implemented in the public sector by more than 65 organisations. Some of the most popular application include:

- Bürokratt – a network of chatbots for different public sector institutions
- AI-driven tool for detecting potential VAT fraud by Estonian Tax and Customs Board
- Computer vision used in distance monitoring by the Environment Agency – for forests, crops, snow, ice, floods, wildlife etc.
- Remote authentication used to perform notarial acts online
- Automatic real-time subtitling on national television
- Automatic transcriptions of court hearings and Parliament sessions. [88]

Public sector AI use cases are outlined on the state AI website kratid.ee.

Despite the robust strategic vision evident in the above recorded policies, Estonia lacks a national AI-specific legal framework. While a bill regulating the effects of algorithmic systems was drafted in 2020 with the aim of identifying necessary legal reforms to accommodate the introduction of AI, this initiative was shelved after the EC announced its intentions to adopt the AI Act. Thus, an opportunity was missed to regulate algorithmic systems that do not necessarily fit into the definition of AI adopted in the AI Act, and potentially offer wider protections. Although AI-focused legislation is missing, laws on

[85] Ministry of Economic Affairs and Communications, Estonia's Digital Agenda 2030. Available at: https://www.mkm.ee/sites/default/files/documents/2022-04/Digiühiskonna%20arengukava_ENG.pdf
[86] Ministry of Economic Affairs and Communications, Andmete ja tehisintellekti valge raamat 2024-2030. Available at: https://www.koda.ee/sites/default/files/content-type/content/2024-02/Valge%20raamat%202024-2030%20%2859%20lk%29.pdf
[87] Ministry of Economic Affairs and Communications, Ministry of Justice, Ministry of Education and Research, Tehisintellekti tegevuskava 2024-2026 (AI Strategy 2024-2026). Available at: https://www.kratid.ee/_files/ugd/7df26f_21000a2dd36c4a66a30eea97563370a3.pdf
[88] Factsheet: AI Strategy. Available at: https://e-estonia.com/wp-content/uploads/factsheet-ai-strategy.pdf

Co-funded by the European Union

the protection of personal data and equal treatment, such as the Personal Data Protection Act, [89] the Equal Treatment Act, [90] and the Gender Equality Act still apply and create obligations for state and private actors. [91] Also relevant is the adoption of an e-State Charter by the National Audit Office and the Chancellor of Justice. The Charter lists the rights of individuals when communicating with public authorities in an e-State, including the right to receive comprehensive information about public services, the right to use one's national e-ID, the right to obtain public services easily and conveniently, the right to receive information about the progress of service provision, the right to know what personal data public institutions have collected and how it is protected, etc. [92]

In addition, the state has made available various support services for the public sector, through the AI Support Toolbox, developed in cooperation between Ministry of Economic Affairs and Communications, the Information System Authority, the Ministry of Justice and the Data Protection Inspectorate to provide support to institutions in launching, carrying out and implementing AI projects.[93] The toolbox includes AI sandboxes, spanning 3-6 months, catering to institutions executing complex AI projects, helping to assess the risks together with an advisory board and to provide practical support in carrying out the project. It also includes expert-led data panels to assess the potential risks of both planned and ongoing projects, which are related to personal data protection or data processing in general. [94]

Complementing these measures are training and capacity building opportunities, including the Digital State Academy, an online platform created by the Ministry of Economic Affairs and Communications in cooperation with a number of public sector partners, that offers free e-courses to anyone interested to learn more the e-State. The platform includes a section on data, with courses on data quality and AI, and a section on equality policies, with courses on equal treatment and gender equality (not specifically related to AI). [95] In addition, the Ministry of Economic Affairs and Communications has an AI specific YouTube channel "Kratid", where trainings relevant to AI are shared. The previous AI Strategy (2022-2023) included trainings to institutions to raise awareness on principles and requirements of human-centred and trustworthy AI, including the potential risks to fundamental rights, as well as guidance materials for assessing and mitigating risks to fundamental rights in the

[89] Personal Data Protection Act (*Isikuandmete kaitse seadus*). Available at: https://www.riigiteataja.ee/en/eli/ee/523012019001/consolide/current
[90] Equal Treatment Act (*Võrdse kohtlemise seadus*). Available at: https://www.riigiteataja.ee/en/eli/507032022003/consolide
[91] Gender Equality Act (*Soolise võrdõiguslikkuse seadus*). Available at: https://www.riigiteataja.ee/en/eli/ee/530102013038/consolide/current
[92] https://www.oiguskantsler.ee/sites/default/files/ESTEriigi%20harta.pdf
[93] Kratid, AI Support Toolbox. Available at: https://www.kratid.ee/en/kratitoe-portfell
[94] Kratid, AI Support Toolbox. Available at: https://www.kratid.ee/en/kratitoe-portfell
[95] Digital State Academy (*Digiriigi Akadeemia*). Available at: https://digiriigiakadeemia.ee/?lang=en

Co-funded by the European Union

development and use of AI. These include data quality guidelines, which stress the importance of reliable and high-quality databases for better decision-making, without, however, specifically addressing the connection between biased data and discrimination. Despite these efforts, the 2024-2026 AI Strategy finds that there is limited awareness and competences in both the public and private sectors about how to deal with possible negative effects in the development and implementation of AI and how to ensure that risks are mitigated.[96]

Finally, the Estonian state is committed to fostering collaboration among different actors from the public and public sectors, as well as research and academia, by creating platforms and opportunities for dialogue and mutual learning. For example, as a part of the European Digital Innovation Hubs (EDIH) network, AI & Robotics Estonia (AIRE) was established in 2022. AIRE brings together academia and industry by helping SMEs develop knowledge-intensive solutions in the field of AI and robotics. AIRE focuses on increasing the capabilities of AI and robotics in the industrial sector. [97] In 2024, the Estonian Centre of Excellence in AI was founded, involving seven research groups from the University of Tartu, four from Tallinn University of Technology and one from Cybernetica AS. The centre aims to advance innovative methodologies for the development of reliable AI systems intended to further AI capabilities in key Estonian sectors, including e-governance, healthcare, business process management, and cybersecurity. [98] Finally, the state makes available open-source AI core components, which are the base components of an application based on AI that all interested parties in the public or private sector are able to reuse without charge and to further develop depending on their own needs. [99]

It is clear that Estonian AI policies mostly concern the adoption of AI in the public sector and largely leave out the private sector. [100] Hence, relevant data paints a very different picture. According to the Digital Society and Economy Index (DESI), 5.2% of Estonian enterprises use some type of AI technology, while the EU average is 8%. [101] The use of AI varies widely across demographic groups – 56% of 18-29-year-olds have used AI-based technologies, but the percentage decreases in every next age group down to 13% for 50-74 age group. Trust in AI-based solutions is the lowest in Estonia of all the Baltic countries: just under one in three Estonians (29%) would trust a service or product based on

[96] Ministry of Economic Affairs and Communications, Ministry of Justice, Ministry of Education and Research, Tehisintellekti tegevuskava 2024-2026 (AI Strategy 2024-2026), p. 37. Available at: https://www.kratid.ee/_files/ugd/7df26f_21000a2dd36c4a66a30eea97563370a3.pdf
[97] AI & Robotics Estonia (AIRE). Available at: https://aire-edih.eu
[98] University of Tartu, „Centre of Excellence in Artificial Intelligence to start work in Estonia", 8 January 2024. Available at: https://ut.ee/en/content/centre-excellence-artificial-intelligence-start-work-estonia
[99] Kratid, Reusable AI components. Available at: https://www.kratid.ee/en/kratijupid
[100] See critique of this approach in relation to the CoE Framework Convention.
[101] European Commission, DESI indicators, Artificial Intelligence, All enterprises. Available at: https://digital-decade-desi.digital-strategy.ec.europa.eu/datasets/desi/charts

Co-funded by the European Union

AI, while nearly half (49%) would not. [102] However, the AI Strategy 2024-2026 sets a goal to expand the availability of AI sandboxes to the private sector. [103] A private sector sandbox service is currently under development. [104]

The field research showed limited interest toward regulating the private sector. According to the interviewed experts, there are very few developers in Estonia who develop AI tools from scratch. Rather they build on top of existing AI, adding layers and criteria. To their knowledge, there are no preventive fundamental rights impact assessment mechanisms in place among private sector developers. One of the interviewed developers was strongly opposed to such impact assessment, as it would slow down the process and make their service more expensive, while violation of rights cannot be completely prevented by the developer, as it depends more on how the tool is used. [105]

The Estonian AI landscape is characterised by an ambivalence between techno-optimism and state-guaranteed fundamental rights protection. One of the main challenges identified in relation to the application of the AI Act is the extensive digitalisation [106] of public administration, which predates the advent of AI, and has not been regulated with the risks of these technologies in mind. Hence, the use of algorithms and AI systems to perform, for example, automated decision-making processes [107] may lead to unforeseen consequences for the fundamental rights of citizens. A 2023 Supreme Court ruling has tackled this issue from the scope of good administration in the area of environmental protection. In a case concerning the automated issuance of a felling permit by the Environmental Board, [108] the Court affirmed that the administrative principles of investigation and caution as well as the obligation to inform the public apply to the administrative procedure and issuance of felling permits regardless of whether the decision is taken by an official or by an automated information system. The Court also emphasised that the administrative body implementing the system is ultimately responsible for the

[102] EY Baltics AI Perception Survey Estonia, March 2024. Available at: https://www.ey.com/en_ee/ai/ey-baltics-ai-perception-survey-estonia-march-2024
[103] Ministry of Economic Affairs and Communications, Ministry of Justice, Ministry of Education and Research, Tehisintellekti tegevuskava 2024-2026 (AI Strategy 2024-2026), p. 40-41. Available at: https://www.kratid.ee/_files/ugd/7df26f_21000a2dd36c4a66a30eea97563370a3.pdf
[104] Interview with Ministry of Justice, 31 July 2024
[105] Interview with an AI developer, 31 October 2024.
[106] P. K. Tupay, M. Mikiver, "The Estonian e-state and challenges of regulating public sector digitalization". *Public Digitalisation in a legal perspective: Status, challenges and opportunities for Nordic-Baltic cooperation.* Nordic Council of Ministers, 13 May 2024, p. 52. Available at: https://pub.norden.org/temanord2024-503/estonia.html
[107] This is particularly prominent in the areas of taxation, environmental protection, and employment: Taxation Act (*Maksukorralduse seadus*), § 46^2 (1). Available at: https://www.riigiteataja.ee/en/eli/505082024003/consolide; Environmental Charges Act (*Keskkonnatasude seadus*), § 33^6 (1). Available at: https://www.riigiteataja.ee/en/eli/ee/513012014001/consolide/current; Unemployment Insurance Act (*Töötuskindlustuse seadus*), § 23 (4). Available at: https://www.riigiteataja.ee/en/eli/519012024006/consolide.
[108] Supreme Court of Estonia, Case No. 3-21-979, 28 September 2023. Available at: https://www.riigikohus.ee/et/lahendid?asjaNr=3-21-979/44

Co-funded by the European Union

legality of any automated administrative decision. In the case of assessment and discretionary decisions, the administrative body may only implement a system that ensures consideration of all important circumstances, and if the available technology does not allow these requirements to be met, a human must participate in the decision-making process. The court stressed that before putting into operation a system that generates automatic administrative decisions, the administrative body must conduct a detailed assessment of the risk of wrong decisions.

The decision leaves open the possibility for automated decision-making without any human intervention at any stage, provided that the AI system takes into account "all important circumstances". However, it is unclear how this can be ensured with respect to all potential cases, which may present an unforeseen degree of complexity that was not anticipated during the design and deployment of the system. This could lead to a violation of the core environmental principle of prevention and to irreversible damage being caused before the relevant decision can be challenged. In addition, the Court requires a risk assessment in terms of the impact of "wrong" decisions made by automated systems and does not seem to consider the potential impact of "right" decisions, which may, for example, require a proportionality assessment requiring human intervention.

The AI fundamental rights impact assessment framework has not yet been developed, although there are plans to address this gap through the EquiTech project[109]. Amendments to the Administrative Procedure Act to regulate automatic administrative acts are in development, meanwhile, certain automatic administrative procedures are already foreseen in specific legal acts – for example, the Tax and Customs Board110, the Environmental Board[111], and the Unemployment Insurance Fund[112] have the right to issue administrative acts and documents in an automated manner without the direct intervention of an official if this is possible considering the circumstances.

Greece

Greece's commitment to artificial intelligence (AI) reflects its broader aspirations to modernise government, improve public services and drive economic growth, while addressing the challenges of

[109] The EU-funded project EquiTech coordinated by the Gender Equality and Equal Treatment Commissioner, Tallinn University of Technology, Ministry of Economic Affairs and the Communications and Ministry of Justice, aims to enhance the capacity of public authorities in addressing risks of discrimination and bias in automated decision-making systems through in-depth research, developing support materials, trainings and a media campaign. More information on the project is available at https://www.volinik.ee/en/interested/ec-projects.html.

[110] Taxation Act (*Maksukorralduse seadus*), § 46^2 (1). Available at: https://www.riigiteataja.ee/en/eli/505082024003/consolide

[111] Environmental Charges Act (*Keskkonnatasude seadus*), § 33^6 (1). Available at: https://www.riigiteataja.ee/en/eli/ee/513012014001/consolide/current

[112] Unemployment Insurance Act (*Töötuskindlustuse seadus*), § 23 (4). Available at: https://www.riigiteataja.ee/en/eli/519012024006/consolide

Co-funded by the European Union

digital transformation. As a member of the European Union (EU), Greece's AI strategy is aligned with EU principles, including the ongoing development of the AI law. However, domestic political, economic and social factors shape the country's AI framework, which seeks to address national priorities such as public administration reform, justice system efficiency and migration management.

Greece's foundational document guiding the adoption of AI is the Digital Transformation Bible (2020-2025), [113] which sets out the country's vision for integrating digital technologies, including AI, into the public and private sectors. The strategy recognises AI as one of the most disruptive technologies of the 21st century, capable of transforming industry and society. Despite this recognition, Greece lags behind other EU member states in the adoption of AI, with limited use by businesses and an underdeveloped national AI strategy. Nevertheless, the Digital Transformation Bible identifies AI as a priority area and outlines key actions, including the creation of centres of excellence, the promotion of innovation and the development of a data governance policy. Central to this vision is the Ministry of Digital Governance, which is leading efforts to ensure that AI benefits all citizens while addressing challenges related to trust, ethics and transparency.

In 2021, Greece participated in the Digital Europe Programme's call for European Digital Innovation Hubs, with the aim of improving digital skills and supporting and enhancing digital innovation, in particular by supporting SMEs and high-tech start-ups in the establishment of artificial intelligence (AI) centres. In December 2021, the Athena Research and Innovation Information Technologies Centre announced the creation of a new independent research unit on artificial intelligence, data science and algorithms, called Archimedes, with a budget of €21 million. Its main objective is to become a gateway to bring together the best scientists in the field of AI, create opportunities and facilitate the transfer of research results to society and the economy. It will serve both basic and applied research, in collaboration with Greek and foreign universities, and will act as a hub providing opportunities for collaboration and synergies between distinguished Greek academics and young scientists in Greece and abroad. It will also foster synergies with the growing start-up ecosystem in Greece to facilitate the transfer of research results to the market. The National Infrastructures for Research and Technology (GRNET) has been providing high-performance computing (HPC) resources through its Advanced Research Information System (ARIS) since 2015. In 2021, a process has been launched to extend the system (budget: €23 million), in close collaboration with the EuroHPC Joint Undertaking. The deployment of the upgraded national HPC/AI centre will also be supported by concrete in-kind contributions from the national HPC centres of several Balkan countries and Cyprus. Greece will be the home of "Pharos", one of the seven AI factories to be

[113] https://digitalstrategy.gov.gr/.

Co-funded by the European Union

developed under EuroHPC. The factory will focus on health, Greek language and culture, and sustainable development. The project, with a total budget of €30 million, will be funded 50% by the EuroHPC Joint Undertaking and 50% by national funds, and is scheduled to start in March 2025, with a total duration of 36 months.

One of the most significant developments in Greece's AI framework is the adoption of Law 4961/2022, which provides the first explicit legal basis for the use of AI in the country. This legislation addresses the use of AI in both the public and private sectors, emphasising transparency, accountability and ethical considerations. It establishes obligations for public bodies to inform stakeholders about the operation of AI systems, conduct algorithmic impact assessments, and ensure compliance with existing data protection laws, such as the General Data Protection Regulation (GDPR). The law also mandates the creation of a national registry for AI systems and establishes oversight mechanisms, including the National Transparency Authority (NTA), to monitor AI system compliance.

The public sector is a key focus of Greece's AI strategy, reflecting its potential to modernise administrative processes and improve service delivery. The first example of AI use in public administration concerns an AI tool developed by the Ministry of Digital Governance, to be used by the Hellenic Cadastre. [114]The tool was deployed in September 2024, and it is designed to accelerate the legal review of property contracts. Prior to its introduction, employees had to read entire contracts themselves, a process that took an average of 30 minutes per contract. The new AI tool analyses the text of the contracts, identifies the type of transaction they concern, verifies that all required information is provided, and formulates a proposal regarding the approval or rejection of the relevant application. The system is designed to significantly cut the time required to process applications, a well-established need that impacts on the right to good administration. However, while the head of department remains responsible for any final decision-making, this system does provide them with ready-made recommendations, which run the risk to be adopted without careful consideration, with implications on property rights and contractual freedom.

Other AI applications, such as the Digital Assistant (a chatbot integrated into the government website gov.gr) and robotic process automation (RPA) technologies, aim to streamline citizen interactions and reduce bureaucratic inefficiencies. The Digital Assistant uses natural language processing to provide answers to citizen queries, a step towards a more citizen-centric government. Similarly, RPA automates repetitive administrative tasks, improving accuracy and freeing up human resources for higher-value tasks. While these applications promise greater efficiency, they also raise concerns about transparency and accountability. For example, relying on automated systems to handle

[114] https://www.ktimatologio.gr/grafeio-tipou/deltia-tipou/1493.

Co-funded by the European Union

citizen requests could obfuscate decision-making processes and limit the ability of individuals to challenge outcomes. The government is also planning AI-based initiatives in relation to the staffing of public services [115] (funded by the Public Investment Program and the Recovery and Resilience Facility of the European Union), as well as in the field of prevention and management of natural and human-induced disasters. [116]

The private public benefit company Hellenic Electricity Distribution Network Operator (HEDNO) is implementing an artificial intelligence (AI) system to combat electricity theft. The technology is used to detect faults in medium voltage lines and enables the prevention of transformer issues. By analysing data from smart meters and the grid, the AI system identifies irregular consumption patterns indicative of unauthorized usage. [117] Although this system does work toward enhancing the efficiency and reliability of the electricity distribution network, it can also be problematic when viewed within the national context. Indeed, electricity theft is observed mostly within severely underprivileged and marginalised communities (e.g., in Roma reservations). If combined with police interventions, this system may lead to profiling, criminalisation, and further alienation of these communities.

In the justice sector, AI is being piloted to improve efficiency and transparency. Two major projects, the Digital Upgrading of Administrative Justice and the Integrated Case Management System, aim to use machine learning to automate the classification of legal documents, identify relevant legislation and facilitate decision-making. These initiatives promise significant benefits, such as reducing backlogs and improving access to justice. But their implementation also raises critical questions about fairness and accountability. The automation of legal processes risks perpetuating biases embedded in training data, while the lack of transparency in AI-generated outputs complicates efforts to ensure judicial independence and procedural fairness. Moreover, the reliance on AI in sensitive areas such as the judiciary underscores the need for robust oversight mechanisms to guard against rights violations.

Migration and asylum management is another area where Greece has embraced AI, often under the lens of national security. The country has deployed AI-enabled border surveillance systems, including facial recognition and anomaly detection technologies, to monitor and control migration flows. While these applications have improved operational capabilities, they have also been criticised for their potential to violate fundamental rights. The AI-driven systems HYPERION and CENTAUR, deployed in Reception and Identification Centres and Closed Controlled Structures for asylum seekers on the island of Samos in Greece, are advanced tools for asylum management and security. HYPERION

[115] https://www.ypes.gr/wp-content/uploads/2022/10/eggr15310-20220929.pdf.
[116] ktpae.gr.
[117] https://www.kathimerini.gr/economy/562060522/techniti-noimosyni-kata-reymatoklopon/.

Co-funded by the European Union

is an asylum management system that collects biometric and personal data, such as fingerprints and facial recognition scans. It links this data to individual cards that are used to manage access to essential services, including healthcare, food and shelter, while tracking movements within and outside reception centres. It operates as a centralised data repository, ostensibly ensuring the efficient distribution of resources to asylum seekers, but also facilitating the tracking of movements, creating a panoptic surveillance environment. CENTAUR is designed to enhance security, using AI algorithms, drones and CCTV cameras to analyse behaviour in real time. Its purpose is to detect aggression or potential escape attempts by classifying behaviour deemed 'suspicious' by its predictive algorithms. The system uses behavioural analytics alongside drone-based perimeter surveillance. Together, these systems represent a paradigm shift in the governance of migration and asylum, moving towards a surveillance-heavy model. The two systems have been the subject of a dispute between the Ministry of Migration and Asylum and the Hellenic Data Protection Authority. In its decision No 13 of 2/4/2024, [118] the DPA fined the Ministry a total of €175,000 for (a) not conducting a full, comprehensive and coherent DPIA during the design phase of the system, thus violating 25 and 35 GDPR; (b) not being transparent and forthcoming with information about the data processing activities of the two systems during the DPA's investigation.

The risks associated with the deployment of AI in Greece highlight the importance of legal and ethical safeguards. Law 4961/2022 addresses some of these concerns by requiring algorithmic impact assessments for public sector AI systems. These assessments evaluate the intended purpose, technical characteristics and potential risks of AI applications, ensuring accountability and trustworthiness. However, gaps remain in the enforcement of these provisions, particularly in the private sector, where transparency obligations are less stringent. For example, companies using AI for employee assessment or recruitment must provide information on the decision parameters, but compliance monitoring is limited, leaving room for potential abuse.

Data protection is another critical area of concern. While GDPR compliance is mandatory, the rapid proliferation of AI systems is challenging regulators' ability to ensure data security and minimise risks. For example, the reliance of digital assistants on open data from public sector websites raises questions about data quality and potential misuse. Similarly, the use of facial recognition technology in border control requires robust safeguards to prevent unauthorised data access and misuse. The National Transparency Authority and the Hellenic Data Protection Authority are tasked with oversight, but their capacity to manage the complexities of AI governance remains limited.

[118] Accessible in Greek at https://www.dpa.gr/el/enimerwtiko/prakseisArxis/aytepaggelti-ereyna-gia-tin-anaptyxi-kai-egkatastasi-ton-programmaton.

Co-funded by the European Union

Public awareness and engagement are also key to Greece's AI strategy. Initiatives to build digital literacy and foster public trust in AI include education programmes and outreach campaigns. However, these efforts are often ad hoc and limited in scope, failing to reach marginalised communities most affected by AI-driven decisions. According to the Digital Economy and Society Index (DESI), Greece has not yet met the challenge of educating its population with the required level of digital skills, despite several recent policies, investments and reforms. Furthermore, only 52.4% of the population had at least basic digital skills in 2023 (EU average 55.5%), indicating no progress since the last data collection in 2021. In 2023, only 43.3% of SMEs had at least a basic level of digital intensity, below the EU average (57.7%). Businesses in Greece also have a low level of adoption of advanced technologies, with 33.5% of firms having adopted AI, cloud or data analytics in 2023, below the EU average of 54.6%. [119]

Despite the challenges, Greece's legal and policy framework for AI is in line with EU principles, emphasising ethical considerations and accountability. The expected adoption of the EU AI Act is expected to provide further clarity, especially for high-risk AI applications such as justice and migration management. The Act's provisions on transparency, bias mitigation and human oversight will strengthen existing safeguards and address gaps in Greece's current framework. However, the country's success in implementing these measures will depend on its ability to strengthen regulatory capacity, foster interdisciplinary cooperation, and prioritise the protection of fundamental rights.

In conclusion, Greece's approach to AI governance reflects its dual aspirations to foster innovation and protect rights. The adoption of Law 4961/2022 is an important step towards a more structured AI framework that addresses transparency, accountability and ethical considerations. However, the practical challenges of enforcing these provisions, coupled with the complexity of AI technologies, highlight the need for continuous adaptation and vigilance.

Malta

Malta's stance on AI deployment and development is shaped by its overall approach to foreign investment as a key driver for continued economic development and social progress. As a country with very limited natural resources, Malta realizes that being technology-savvy [120] is key to winning regulation competition with other EU member states and other countries. The public perception of AI in Malta is very positive. It has also received a very positive coverage in the general press with

[119] https://digital-strategy.ec.europa.eu/en/factpages/greece-2024-digital-decade-country-report.
[120] The same approach was adopted in respect of other similar areas, such as intellectual property ('the IP hub'), blockchain (the 'blockchain island') and cryptocurrencies.

Co-funded by the European Union

references made to the possibilities AI presents for medicine and public health[121], economic growth[122] and societal benefits.[123] The discussions of the risks posed by AI, including its fundamental rights implications, are rare and superficial in the general press,[124] while more sophisticated and in-depth discussions are confined to specialized circles. In 2019, the government rolled out the 'Strategy and Vision for Artificial Intelligence in Malta 2030', [125] drafted by the Malta.AI Taskforce [126] commissioned by the Maltese government. The main objective of the national strategy is to make Malta a global leader in the area of AI, thus gaining a strategic competitive advantage in the global economy.

The Maltese AI Strategy relies on three strategic pillars: investment and innovation; public sector adoption; and private sector integration. Accompanying these pillars are enablers such as education, ethical guidelines, and infrastructure development. Despite its ambitious proclamations, however, the operationalisation of the policies and measures it envisions is covered by obscurity or remains on paper. Specifically, although some sources [127] mention total of 72 AI actions undertaken in the public sector, jointly developed with relevant stakeholders who have committed to taking responsibility for achieving the common targets, no details of those actions are publicly available. Similarly, the fate of the six AI projects mentioned in the strategy is unknown. One of these – traffic management – is reportedly paused due to the resignation of a key team member. Increasing concerns over opacity and oversight, the latest updates on the Malta.AI Taskforce website are dated March 27, 2019.[128] Finally, measures to digitally upskill workers and mitigate job loss due to AI are still in the planning stage. On a positive note, AI awareness building among students, parents, and teachers is progressing, with various activities carried out at different levels of education, [129] aided by the government's scholarship scheme on AI.[130]

---

[121] See https://timesofmalta.com/articles/view/artificial-intelligence-in-the-medical-field.891190; https://timesofmalta.com/articles/view/artificial-intelligence-may-be-pandemic-lifesaver-one-day.780998; https://timesofmalta.com/articles/view/want-to-control-your-bed-with-just-your-thoughts-theres-an-app-for.906346?fbclid=IwAR0_ILHcgtF8I803ELpOuEiECgoEA1a6986OFzPWELYtfMIsMquYn0hClNc

[122] See https://timesofmalta.com/articles/view/where-next-for-malta-and-ai.724963; https://timesofmalta.com/articles/view/saudi-to-invest-20-billion-in-ai-by-2030.832969; https://timesofmalta.com/articles/view/game-technology-firm-with-malta-university-input-charts-growth.845016

[123] See https://timesofmalta.com/articles/view/ai-to-control-traffic-in-pilot-project.739657

[124] See, e.g., https://timesofmalta.com/articles/view/understanding-maltas-artificial-intelligence-framework.747060

[125] Ibid.

[126] See https://malta.ai/the-team/

[127] https://digital-skills-jobs.europa.eu/en/actions/national-initiatives/national-strategies/malta-strategy-and-vision-artificial-intelligence

[128] https://malta.ai/news/

[129] See, e.g., https://eskills.org.mt/training-offer/?search_term=&field_digital_technology%5B%5D=http%3A%2F%2Fdata.europa.eu%2Fuxp%2F3030&field_is_free=&field_training_start_date.

[130] See https://businessnow.mt/second-round-for-ai-postgraduate-scholarship-fund-announced-with-wider-criteria/.

Co-funded by the European Union

On the flip side, Malta's private sector has been actively deploying AI in various areas, most prominently in:

- **Customer service.** AI-powered chatbots and virtual assistants are increasingly used to handle routine inquiries, provide quick responses, and improve overall customer satisfaction. According to the field research, AI-driven customer support is becoming prevalent in banking, retail, and other industries.
- **Marketing.** AI applications are used for automated split testing and optimisation, dynamic user-specific content presentation, natural language processing (NLP), etc.
- **Finance.** In banking, insurance and investment, AI applications are used for multiple functions, such as onboarding, fraud detection, compliance, and accounting.
- **iGaming.** In iGaming, including online gambling and betting, AI technologies are frequently incorporated into solutions for detecting and reducing fraud, enhance marketing effectiveness and augment customer service interactions and customer experience functions.
- **Software development.** A whole range of AI applications are being used in Malta for coding, testing and bug fixing (e.g., CodeGPT, Claude, Codi).
- **Professional services.** Maltese lawyers, accountants, auditors and other professionals are increasingly relying on AI applications for NLP, project management (e.g., make.com), and research (e.g., Luminance).
- **Manufacturing.** In the advanced manufacturing, such as electronics [131] and aviation maintenance, repair and overhaul industries, AI-driven solutions are being deployed for condition monitoring and predictive maintenance activities. The solutions draw on the vast amount of data that modern aircraft, vehicles and other machinery generate.
- **Tourism.** Malta's tourism industry benefits from AI applications, including personalized travel recommendations, dynamic pricing for accommodations, and chatbots that assist tourists with information about attractions, events, and local services.

In addition to the AI Strategy, Malta has also adopted a national framework for the development of ethical AI: "Malta – Towards Trustworthy AI".[132] This framework is the product of public consultation and aims to set out guiding principles and governance practices for trustworthy AI in Malta and

[131] E.g., in the STMicroelectronics that has run a production line in Malta since 1981
[132] See https://malta.ai/wp-content/uploads/2019/08/Malta_Towards_Ethical_and_Trustworthy_AI.pdf

Co-funded by the European Union

beyond, mirroring those established in the AI HLEG Ethics Guidelines for Trustworthy AI: human autonomy, prevention of harm, fairness, and explicability.[133]

Information on the methods or tools used in both the public and private sectors to assess the impact of AI systems on individual rights is not publicly available. Accordingly, the mechanism by which decisions to deploy AI are made in both sectors, who is involved in making those decisions, what data they examine, and whether there are tools to evaluate or assess relevant decisions also remain unknown to the public. Similarly, there is no information on whether AI developers and users assess the fundamental rights implications arising from these technologies and how they perform this assessment. Nevertheless, the Maltese government has set out a Technological Assurance Sandbox (TAS), offered by the Malta Digital Innovation Authority (MIDA). MDIA-TAS is intended to guide innovators throughout a residency of maximum four years, as they align their technological solution with established Control Objectives based on international standards, including legal and ethical ones. At the end of each phase of this alignment, an independent third-party technical assessment is conducted by a selected MDIA-authorised Systems Auditor (or an MDIA-recognised Technical Expert), until compliance is achieved with all Control Objectives and/or proposed Milestones. After a defined number of assessments, the Applicant can obtain the full MDIA certification, indicating that their AI solution provides technological assurances for various stakeholders, including users and investors. [134] Unfortunately, it is not publicly known if any company has exercised this option to date.

Besides the directly applicable EU AI Act, there is no AI-specific legislation in Malta. MIDA is the mandated oversight body on AI. Its mission is "to promote consistent principles for the development of visions, skills and other qualities relating to innovative technology, and to exercise regulatory functions regarding innovative technology and related services and to provide for matters ancillary thereto or connected therewith". In addition, the Malta Information Technology Agency (MITA) is responsible for providing ICT infrastructure, systems and services to the government. MITA is supposed to follow the ethical guidelines provided by the "Malta – Towards Trustworthy AI". However, it is unclear whether it does so or not.

Compounding the lack of transparency and a clear regulatory framework (with the exception of the – directly applicable – AI Act), is the limited awareness of citizens and the civil society with regards to the ethical and fundamental rights risks created by AI. CSOs are primarily concerned with other systemic issues, such as the rampant corruption, organised crime, intimidation of civil society activists

[133] See https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

[134] See https://www.mdia.gov.mt/technology-assurance-sandbox/.

Co-funded by the European Union

and journalists, etc., which they consider more pressing. [135] AI falls low on their list of priorities. A research participant mentioned that, "until a big scandal happens", AI is not high on the public agenda.

The Netherlands

The Netherlands' approach to AI governance has been significantly influenced its previous experiences with the use of algorithmic systems and notably dealing with the aftermath of harms caused by algorithmic systems in the so-called child benefits scandal. [136] This chapter critically examines the Netherlands' legal and policy framework for AI, emphasizing its strategic initiatives, regulatory measures, and the challenges associated with protecting fundamental rights in a rapidly evolving technological landscape.

The foundation of AI policy in the Netherlands lies in its **National AI Strategy**, adopted in 2019. This strategy focuses on accelerating AI innovation while embedding ethical considerations, trust, and public values at the core of its implementation. Central to the strategy are three pillars: strengthening AI research and innovation, fostering talent and skills, and promoting the responsible use of AI in public and private sectors. The strategy's emphasis on public values is significant, as it reflects a recognition of the societal risks posed by AI, including potential infringements on privacy, non-discrimination, and access to justice.

The governance of AI in the Netherlands is distributed across multiple institutions, with the Ministry of Economic Affairs coordinating the overall strategy. Other ministries, including Justice and Security, Social Affairs, and the Interior, oversee AI applications within their respective domains. This decentralized approach allows for tailored oversight but also creates challenges in ensuring consistency and accountability across sectors. The Dutch Data Protection Authority (Autoriteit Persoonsgegevens) plays a critical role in safeguarding privacy, while the Netherlands Institute for Human Rights focuses on issues of equality and non-discrimination.

One of the most impactful moments in shaping the Netherlands' AI policies was the **childcare benefits scandal**. Between 2013 and 2019, the Dutch Tax Authority used an algorithm to detect fraud in childcare benefits applications, relying on factors such as dual nationality as a risk indicator. This algorithmic profiling led to the wrongful accusation of fraud against approximately 26,000 parents, 76% of whom were from ethnic minority groups. The scandal resulted in severe personal and financial consequences for affected families, including the loss of housing, employment, and, in extreme cases, custody of children. The public outrage and political fallout from this scandal, which culminated in the

[135] See, e.g., https://repubblika.org/.
[136] For elaborate information on the algorithm and its impacts, see this report of Amnesty Netherlands: https://www.amnesty.org/en/latest/news/2021/10/xenophobic-machines-dutch-child-benefit-scandal/

Co-funded by the European Union

resignation of the Dutch cabinet in 2021, underscored the dangers of poorly governed AI systems and catalysed reforms aimed at protecting fundamental rights.

The government's response to the scandal included the introduction of tools designed to ensure the ethical deployment of AI. Among these, the **Fundamental Rights and Algorithmic Impact Assessment (FRAIA, or IAMA in Dutch)** [137] stands out as a mechanism to assess the risks posed by algorithms to fundamental rights. Although not yet mandatory, IAMA has been piloted by public institutions and offers a structured approach to evaluating the ethical and legal implications of AI systems. Complementing this is the **Algorithm Register**, which encourages transparency by requiring public sector institutions to disclose information about their use of impactful algorithms. However, the voluntary nature of these tools and inconsistencies in their implementation limit their effectiveness.

Another cornerstone of the Netherlands' AI governance framework is the **Algorithm Implementation Framework (IKA)**, developed by the Ministry of the Interior. This "living document" integrates various assessment and accountability standards, drawing on resources such as the IAMA tool and the Netherlands Court of Audit's methodology for algorithmic risk evaluation. The framework aims to provide clear benchmarks for responsible AI use, addressing issues such as bias, accountability, and human oversight. Despite its comprehensive design, the IKA remains underutilized, with many public institutions struggling to operationalize its recommendations effectively.

The Netherlands' private sector also plays a significant role in AI adoption, particularly in industries such as finance, information technology, and research. Companies use AI for tasks ranging from credit risk assessment to customer service optimization. However, fundamental rights considerations are often secondary to business objectives. While some companies incorporate ethical frameworks and human rights principles into their operations, many limit their compliance efforts to meeting the minimum legal requirements under the General Data Protection Regulation (GDPR) and forthcoming EU AI Act. This compliance-focused approach highlights a broader challenge in aligning corporate interests with societal values.

Despite these advancements, the Netherlands faces persistent challenges in its AI governance. One notable issue is weak enforcement and oversight. Civil society organizations and legal experts have criticized the Dutch Data Protection Authority for its perceived inaction, particularly its failure to investigate and penalize violations of fundamental rights. This undermines public trust in AI governance and raises questions about the effectiveness of existing accountability mechanisms. Another factor is that a lot of the available information and experience is siloed between - and even within – various

[137] https://www.government.nl/documents/reports/2021/07/31/impact-assessment-fundamental-rights-and-algorithms.

Co-funded by the European Union

actors working in this field. There is very little knowledge transfer, leading to everyone involved constantly re-inventing the wheel instead of learning from each other.

Public awareness of AI's risks has increased significantly in the wake of the childcare benefits scandal and subsequent investigations by journalists and advocacy groups. However, this awareness is uneven, with a stronger focus on privacy concerns than on broader issues such as indirect discrimination or access to justice. Civil society organizations have been instrumental in raising awareness and advocating for stronger protections, but their efforts are often siloed, limiting their reach and impact. The government has sought to address this gap through public discussions and educational initiatives, yet these efforts remain sporadic and lack a coherent strategy.

The Netherlands has also grappled with the technical and practical challenges of embedding fundamental rights considerations into AI systems. Engineers and data scientists often lack formal training in ethics or human rights, relying instead on ad hoc methods to identify and mitigate risks. This gap underscores the need for interdisciplinary approaches that integrate legal, technical, and ethical expertise. Initiatives such as the creation of ethical advisory committees within municipalities represent a step forward, but these efforts are fragmented and lack standardized practices.

The Dutch judiciary has begun to engage with AI-related issues, most notably through the **System Risk Indication (SyRI)** case. In 2020, The Hague District Court ruled that the legislation underpinning SyRI, a tool used for fraud detection in social benefits, violated Article 8 of the European Convention on Human Rights (ECHR). The court emphasized the lack of transparency and proportionality in SyRI's operations, setting a precedent for the judicial scrutiny of algorithmic systems. Despite this landmark decision, there is limited information on the extent to which judges receive training on AI, raising concerns about the judiciary's preparedness to handle complex algorithmic cases.

Looking ahead, the Netherlands' adoption of the EU AI Act is expected to provide a more robust regulatory framework, particularly for high-risk AI applications. The Act's requirements for transparency, accountability, and bias mitigation align with the country's existing initiatives and offer an opportunity to strengthen enforcement and oversight. However, the success of these measures will depend on their implementation at the national level, particularly in addressing gaps in institutional capacity and public engagement.

In conclusion, the Netherlands' legal and policy framework for AI reflects a commitment to balancing innovation with fundamental rights. While the country has made significant progress in developing tools and frameworks to govern AI, challenges remain in enforcement, transparency, and public awareness. Addressing these issues will require a concerted effort to build institutional capacity, foster interdisciplinary collaboration, and engage with diverse stakeholders. By doing so, the

Co-funded by
the European Union

Netherlands can ensure that AI serves as a force for social good while upholding the principles of fairness, accountability, and human dignity.

Co-funded by
the European Union

AI and Fundamental Rights in practice

While educated guesses about the impact of AI technologies in the future are both possible and useful, it is important to examine current applications of AI to avoid venturing into the realm of speculation. This section will highlight major fundamental rights concerns, stemming from the deployment and use of AI in certain crucial sectors, citing specific instances where rights – as enshrined in the EU Charter – are jeopardized by AI systems that are currently in operation. The taxonomy of capabilities followed here is presented below. [138]

However, as noted in the Introduction, AI also presents opportunities to promote the exercise of fundamental rights, if designed carefully and with the appropriate safeguards in place. Examples of such human-centric models are mentioned in the second part of this chapter.

Finally, the chapter presents mitigating measures which are necessary to ensure that AI is used responsibly and respects fundamental rights and ethical principles.

AI Capabilities Taxonomy

1. **Computer Vision**
   - Image segmentation
   - Object detection and tracking
   - Image classification
   - Emotion recognition
   - 3D reconstruction
2. **Computer Audition**
   - Speech to text
   - Musical knowledge
   - Sound similarity assessment
   - Source separation
   - Audio-based sentiment analysis
3. **Computer Linguistics**
   - Translation
   - Text classification
   - Sentiment analysis
   - Entity recognition
   - Relation extraction
   - Conversational systems
4. **Robotics**
   - Robot motion planning
   - HD mapping and localization
   - Control optimization
   - Collaborative robotics / human robot interaction
   - Advanced drones
   - Mobile robotics
   - User-adaptive control automation

[138] Taken from: *Creation of a Taxonomy for the European Ecosystem*, European Institute of Innovation and Technology (EIT), Accessible at https://eit.europa.eu/sites/default/files/creation_of_a_taxonomy_for_the_european_ai_ecosystem_final.pdf.

Co-funded by the European Union

5. **Forecasting**
    - Time series forecasting
    - Dependency-based forecasting
6. **Discovery**
    - Segmentation and clustering
    - Anomaly / outlier detection
    - Correlation analysis
    - Causal inference
    - Association analysis
7. **Planning**
    - Cooperative multi-agent systems
    - Policy development / Strategic agents
    - Logistics planning
    - Planning and scheduling
8. **Creation**
    - Audio generation
    - Image generation / manipulation
    - Style transfer
    - Text generation / summarization
    - AI-augmented engineering

Co-funded by
the European Union

Risks

Table 3: AI in Law Enforcement

Ai in law enforcement	
Relevant AI capabilities	• Computer vision (facial and emotion recognition, detection and tracking, etc.) • Computer audition (sound assessment, audio-based sentiment analysis, etc.) • Computer linguistics • Risk assessment tools to assess likelihood of recidivism • Linguistic analysis algorithms – to detect false information • Acoustic threat detection algorithms – to detect potential violent threats
Charter rights potentially affected	• Article 2, Right to life • Article 3, Right to the integrity of the person • Article 6, Right to liberty and security • Article 7, Respect for private and family life • Article 8, Protection of personal data • Article 21, Non-discrimination • Article 41, Right to good administration • Article 47, Right to an effective remedy and to a fair trial • Article 48, Presumption of innocence and right of defence
Examples of AI systems currently in use	• **Harm Assessment Risk Tool (HART) – HART** [139] is an AI system trained by University of Cambridge criminologists and used by UK police to predict the risk of a suspect reoffending. It uses 104,000 histories of people previously arrested and processed and looks at vast numbers of combinations of 'predictor values' based on the "random forests" method. The majority of these values focus on the suspect's offending history, as well as age, gender and geographical area.

[139] https://www.cam.ac.uk/research/features/helping-police-make-custody-decisions-using-artificial-intelligence.

Co-funded by the European Union

	• **Facewatch** [140] is a privately developed and owned facial recognition system, deployed in spaces offering retail services in the UK, to instantly identify past offenders of shoplifting and alert security personnel. • **VeriPol** [141] is an AI tool used by the Spanish National Police to detect to detect false police reports through a combination of Natural Language Processing and machine learning classification algorithms. It is primarily used in cases involving low level crimes, although the possibility to develop additional functionalities to detect other forms of crime is being considered. • **ShotSpotter** [142] is a gunshot-detection system designed to identify and locate gunfire in real time. It uses a network of acoustic sensors placed in urban areas to detect the sound of gunfire. When the system detects a potential gunshot, it analyses the sound and sends an alert to law enforcement, providing information such as the location of the incident, the number of shots fired and the time it occurred. • **Clearview AI** [143] acts as a search engine of publicly available images, including through social media – now more than 50 billion – to support investigative and identification processes trough faceprints — unique biometric identifiers akin to a fingerprint or DNA profile. The Clearview database has been used by private companies, police, and federal agencies in the USA. Clearview purports to require of its customers human assessment of its results, although it is unclear how this "obligation" can be monitored and enforced. A number of European regulatory authorities, including the French SA and the Greek DPA, have issued steep fines (to the tune of € 20 000 000) against Clearview, for violations of data protection legislation.

[140] https://www.facewatch.co.uk/.
[141] https://ai-watch.github.io/AI-watch-T6-X/service/90122.html.
[142] https://www.soundthinking.com/law-enforcement/leading-gunshot-detection-system/.
[143] https://www.clearview.ai/.

Co-funded by the European Union

	• **The Crime Anticipation System (CAS)** [144] used by the Dutch police, is a geographic crime prediction algorithm to assess crime rates in specific areas. Using anonymised historical police records, aggregated socio-economic data about the area and street location data, the system calculates how many incidents of a particular type of crime occurred in or around this area in a 12-week period and how many known suspects of that type of crime lived in the vicinity of the compartment during this period. • **The Amsterdam Top400 and Top600 lists** [145] are two programmes run by the Municipality of Amsterdam that aim to reduce the number of high-impact crimes, such as robberies, burglaries, serious violence, murder, drug-trafficking and violent crimes, by intervening in the lives of individuals identified as risky. The Top600 programme identifies known high impact offenders who are most at risk of reoffending. The Top400 programme introduced subsequently targets young people between 12 and 23 years old, often siblings (usually brothers) of people listed on the Top600, who have not committed serious offences but who – according to the municipality – need closer observation. • **Predictive Policing algorithms used by the French police.** The French police use several predictive policing software systems, most of which encompass AI elements. Examples include **RTM (*Risk Terrain Modelling*)**, a "situational prevention" software program used by the Paris Police Prefecture to target intervention zones based on "environmental" data (presence of schools, shops, metro stations, etc.); **M-Pulse**, previously named Big Data of Public Tranquility, developed by the city of Marseille in partnership

[144] https://algoritmes.overheid.nl/en/algoritme/81228922
[145] https://algoritmes.overheid.nl/en/algoritme/top-400600-municipality-of-amsterdam/75856898#verantwoordGebruik

Co-funded by the European Union

	with the company Engie Solutions to assess the suitability of municipal police deployments in urban public space; **Smart Police**, an application that includes a "predictive" module aimed at addressing delinquency to improve public safety, developed by French startup Edicia and sold this to over 350 municipal forces.
Risks of fundamental rights' violations	The risks of AI use in the context of law enforcement are many and touch on many different rights. Risks to **liberty and security** of the person are always at stake when it comes to potential arrests and the subsequent implication in criminal cases. When arrests are performed on the basis of contested science or inaccurate or incomplete data, the risk of potential violations increases significantly. Moreover, depending also on the police culture in each individual country, inaccurate information, either on the identity of a suspect or on the type of crime being committed may lead to excessive use of force and violations of the rights to **life** and to **the integrity of the person.** For example, research reveals that only 16,57% of the alerts produced by ShotSpotter have led to confirmed incidents of gunfire, increasing the potential for a lethal response to innocuous cases. This is especially true when combined with the confirmed **racial biases** of this system based on the demographics of the area when the supposed gunshot took place. [146] The lack of accuracy and discriminatory outcomes of ShotSpotter have led to various initiatives calling for its abolishment. [147] **Discrimination** is indeed another major concern, which also occurs in less glaring examples. Detecting and preventing all underlying societal biases in training data is a near impossibility, leaving law enforcement, and in particular predictive policing, open to inherent and grave risks. Faulty police investigations and arbitrary arrests can have ramifications spanning the entirety of the criminal proceedings, potentially undermining **the rights to an effective remedy and to a fair trial** as well as the **presumption of innocence and right of**

[146] https://www.documentcloud.org/documents/25444987-brooklyn-defenders-shotspotter-report/.
[147] https://stopshotspotter.com/.

Co-funded by the European Union

	defence. In addition, based on the "fruit of the poisonous tree" doctrine, illegally obtained primary evidence can lead to the negation of the entire proceedings, wasting court resources, undermining trust to the criminal justice system, and delaying the administration of justice, while also – potentially – creating security risks. In addition, the use of proprietary algorithms and software used in many privately developed AI systems creates concerns over transparency and accountability and runs contrary to the right to **good administration**. Last but not least, **privacy and personal data protection** can be severely undermined by facial recognition systems, detection and tracking and other similar technologies. For example, the Clearview company has been successfully sued by the American Civil Liberties Union (ACLU) for violations of privacy. [148] It should be noted that many of the described systems are contrary to the AI Act, which proscribes the use of systems that: a) make risk assessments on the likelihood of a natural person committing a criminal offence, based solely on profiling or on their personality traits and characteristics, without human involvement, b) create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage, c) 'real-time' remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement, with the – extensive – caveat of Article 5 (1) (h) AI Act.

Table 4: AI in Justice

Ai in Justice	
Relevant AI capabilities	• Computer audition (e.g., speech to text applications) • Computer linguistics (translation, text classification, categorisation, conversational systems, e.g., chatbots, etc.)

[148] https://www.aclu.org/press-releases/big-win-settlement-ensures-clearview-ai-complies-with-groundbreaking-illinois.

Co-funded by the European Union

	• Dependency-based forecasting (prediction of likely outcomes) • Correlation analysis and causal inference (e.g., in the assessment of flight risk, likelihood of recidivism, etc.) • Logistics, planning, and scheduling (e.g., office management, docket entries, case prioritisation, etc.) • Text generation / summarization
Charter rights potentially affected	• Article 3, Right to the integrity of the person • Article 6, Right to liberty and security • Article 20, Equality before the law • Article 21, Non-discrimination • Article 23, Equality between men and women • Article 47, Right to an effective remedy and to a fair trial • Article 48, Presumption of innocence and right of defence • Article 49, Principles of legality and proportionality of criminal offences and penalties.
Examples of AI systems currently in use	• **Correctional Offender Management Profiling for Alternative Sanctions (COMPAS)** [149] is a case management and decision support tool to assess the likelihood of recidivism. It is deployed and used in the US criminal Justice system and assesses factors such as criminal history, substance abuse, and social environment, generating risk scores for pre-trial recidivism (flight risk, risk of committing new felonies), general recidivism, and violent recidivism. These scores are intended to aid probation officers, judges, and other professionals in making informed decisions about bail, sentencing, parole, and intervention programs. • **Spain** has deployed a number of systems in its justice sector, intended for use by the judiciary, court administrators, lawyers and the general public, to facilitate research, decision-making, anonymisation and pseudonymisation,

[149] https://www.documentcloud.org/documents/25444987-brooklyn-defenders-shotspotter-report/. https://www.researchgate.net/publication/321528262_Correctional_Offender_Management_Profiles_for_Alternative_Sanctions_COMPAS.

Co-funded by the European Union

	prioritisation of cases, filling, transcription, translation, and provision of legal information. **Italy** has introduced similar systems, with the addition of a system that predicts the likely outcomes of cases. **Germany** has been deploying AI systems in the area of justice since 2008, most prominently for research, filling, transcription and translation. These systems are developed by the State and academia. [150] • **Lexebra** [151] is an AI-based tool that analyses Bulgarian case law, with the aim to save time and resources for legal practitioners by summarising key arguments of civil, commercial, and criminal case law. It draws on data from over 250,000 cases, focusing on decisions of the Supreme Court of Cassation, particularly those that harmonise the case law or develop the law, interpreting the exact meaning of the legal provisions. • A variety of **legal chatbots** have been deployed in the private sector in recent years. The Estonia MoJ has commissioned the legal tech startup **HUGO.legal** to create software that offers affordable legal services to Estonian residents through a sharing economy-based marketplace of lawyers who provide legal services up to 3 times more affordable than the average hourly rate on the market, as well as by directly offering legal counsel through lawbots-as-a-service. [152] In addition to the general public, the tool is also addressed to lawyers and provides them with services related to client management, work automation, and a lawbot-assistant to minimize their non-billable hours. The **LexisNexis** [153] database has developed several AI components (Lexis Create+, Lexis+, Nexis+, Nexis Data+) for legal research, summarised data, and drafting of documents through

150 https://public.tableau.com/app/profile/cepej/viz/ResourceCentreCyberjusticeandAIFR/AITOOLSINITIATIVESREPORT.
151 https://www.lexebra.com/.
152 https://arcticstartup.com/estonia-orders-e3-5-million-worth-of-legal-services-from-a-startup-bot/.
153 https://www.lexisnexis.com/en-us.

Co-funded by the European Union

	generative AI, drawing on its vast library and available in the jurisdictions it serves. The legal office management software **Clio** has also ventured into AI territory with Clio Duo, [154] which not only helps with case management but also extracts and summarises information derived from the archive of the law firm using it. OpenAI has developed **CoCounsel**, [155] an LLM designed to perform routine and sophisticated legal tasks, including legal research memo drafting, deposition preparation, document review, and others, upon launching. The tool is currently on beta testing but according to the press release issued, it aims, among others, to improve access to justice for marginalised communities by decreasing the cost of legal services.
Risks of fundamental rights' violations	At the outset, tools for **planning and organisation**, as well as tools to assist with the performance of ancillary tasks (e.g., **transcription, translation**, etc.) do not appear to present any significant risks for fundamental rights. In fact, the introduction and use of such tools in justice sector may be beneficial to the rights concerned (see above), particularly by contributing toward the speedy resolution of disputes and criminal cases, a major component of **fair trials**. The involvement of the state and/or academia in the development of AI for courts is likely to add a layer of additional protection which may not be present in for-profit applications. A key caveat here is the careful development of said tools, including in relation to data security (courts are a likely target of cyber-attacks and ransomware and the increased digitalization of justice may create additional risks in this respect). Moreover, mistakes in the prioritization of cases based, for example, on their likely outcome, may have the opposite effect, and delay the resolution of urgent cases to the detriment of justice and fairness. Human intervention can be key to prevent or rectify these situations.

[154] https://www.clio.com/blog/clio-duo/.
[155] https://www.cocounsel.ai/.

Co-funded by the European Union

	AI tools introduced to support **decision-making** (e.g., through the assessment of risks of flight, recidivism, etc.) give rise to more concerns, if not carefully regulated. **COMPAS**, for example, has become the target of criticism due to its opaqueness. It essentially operates as a black box, while its developer has refused to share information with interested parties on how the software weighs particular input variables and how these inputs are calculated for the final risk score, citing the need to protect their trade secrets. [156] In the case of defendant Eric Loomis, [157] who alleged that COMPAS illegitimately considered his gender as a factor for potential recidivism, this resulted to a six-year prison sentence without a fully reasoned, public decision, and had a clear impact on his right to **equality of arms** and – possibly – to **non-discrimination**. **Bias** is another major issue when it comes to AI in justice, whenever AI systems are involved in decision-making as well as summarisation, extraction of key arguments which may require social context, etc. Societal biases present in training data or inserted into the design of the relevant systems, are likely to be replicated in the results produced. On the other hand, it has also been argued that AI can help mitigate societal biases and misconceptions inevitably influencing decision-making in the area of justice. [158] Last but not least, although certain lawbots or similar applications may increase **access to a lawyer**, especially in marginalised communities, the risks for **quality legal representation** and, ultimately, **equal access the justice system** are considerable. Lawyers using these tools risk becoming complacent and overly dependent on them, potentially overlooking crucial details which may require a "human eye". In addition, access to a lawyer can also

[156] See Taylor R. Moore, *Trade Secrets and Algorithms as Barriers to Social Justice*, Centre for Democracy and Technology (2017), accessible at https://cdt.org/wp-content/uploads/2017/08/2017-07-31-Trade-Secret-Algorithms-as-Barriers-to-Social-Justice.pdf.
[157] AI: Tool or obstacle for delivering justice, accessible at https://tecscience.tec.mx/en/human-social/justice-with-artificial-intelligence/.
[158] See, e.g., https://justice-trends.press/ai-for-justice-tackling-racial-bias-in-the-criminal-justice-system/.

Co-funded by the European Union

	be compromised where outcome predictors are being used and lawyers cherry-pick the cases they are willing to take on. At the same time, it is also true that AI can process a vast number of documents in a fraction of the time a human requires to do so and without the inevitable fatigue, ultimately reducing the probability of error and allowing lawyers to focus more on strategy, argumentation, etc. Ultimately, human intervention is key to mitigate these risks.

Table 5: AI, Public Participation, and Citizens rights

AI, Public Participation, and Citizens rights	
Relevant AI capabilities	• Sentiment analysis • Relation extraction • Forecasting • Correlation and association analysis
Charter rights potentially affected	• Article 7, Respect for private and family life • Article 8, Protection of personal data • Article 10, Freedom of thought, conscience and religion • Article 11, Freedom of expression and information • Article 21, Non-discrimination • Article 25, The rights of the elderly • Article 26, Integration of persons with disabilities • Article 39, Right to vote and to stand as a candidate at elections to the European Parliament • Article 40, Right to vote and to stand as a candidate at municipal elections • Article 47, Right to an effective remedy and to a fair trial
Examples of AI systems currently in use	• Cambridge Analytica, a British consulting firm, was famously implicated in a major political scandal, unveiled in 2018. The company illicitly acquired data of up to 87 million Facebook

Co-funded by the European Union

	users, [159] to use in the political campaigns of its clients. The data was used for microtargeted advertising aimed at influencing voter behaviour by showing them content that would resonate with their personality profiles. AI played a crucial role in the data analysis, profiling, and running of microtargeting campaigns. [160] According to its former CEO, Alexander Nix, CA was involved in 44 U.S. political races in 2014; [161] performed data analysis services for Ted Cruz's presidential campaign in 2015; [162] and worked for Donald Trump's presidential campaign and "Leave EU" (one of the organisations campaigning for the Brexit referendum) in 2016. [163] Facebook had to settle a lawsuit for failing to protect the personal data of its users, for $725m. [164]
Risks of fundamental rights' violations	The example of the Cambridge Analytica scandal highlights the threat AI may pose το **fundamental democratic principles** and **self-determination**, if left unregulated. The scandal not only violated the **privacy and personal data** of millions, but arguably also played a major role in crucial moments of **public participation**, that shaped the future of individual countries and the world. Voter manipulation impacts on both the **right to vote and to stand candidate** in elections, facilitating access to political power for those that can afford to pay for it and distorting the very essence of democracy. At the same time, the fake news and biased rhetoric often relied on in this type of campaigning, restrict the **right to information** and the **freedom of conscience** by interfering in the cognitive process of formulating political opinions and decisions through psychological

[159] https://www.nytimes.com/2018/04/04/technology/mark-zuckerberg-testify-congress.html.
[160] https://www.politico.eu/newsletter/ai-decoded/politico-ai-decoded-how-cambridge-analytica-used-ai-no-google-didnt-call-for-a-ban-on-face-recognition-restricting-ai-exports/.
[161] https://web.archive.org/web/20170928200643/https://www.washingtonpost.com/web/20170928200643/https://www.washingtonpost.com/politics/cruz-campaign-paid-750000-to-psychographic-profiling-company/2015/10/19/6c83e508-743f-11e5-9cbb-790369643cf9_story.html?utm_term=.32903dcd5bff.
[162] Ibid.
[163] https://www.nytimes.com/2017/03/06/us/politics/cambridge-analytica.html; https://www.theguardian.com/technology/2017/mar/04/cambridge-analytics-data-brexit-trump.
[164] https://www.bbc.com/news/technology-64075067.

Co-funded by the European Union

	profiling and false information. This process **disproportionately harms the most vulnerable**, including people of precarious socioeconomic status, people with limited digital literacy, the elderly and persons with disabilities.

Table 6: AI in Social Security and Welfare

AI in Social Security and Welfare	
Relevant AI capabilities	• Forecasting • Correlation and association analysis
Charter rights potentially affected	• Article 7, Respect for private and family life • Article 8, Protection of personal data • Article 20, Equality before the law • Article 21, Non-discrimination • Article 22, Cultural, religious and linguistic diversity • Article 23, Equality between men and women • Article 24, The rights of the child • Article 25, The rights of the elderly • Article 26, Integration of persons with disabilities • Article 29, Right of access to placement services • Article 34, Social security and social assistance • Article 36, Access to services of general economic interest • Principle of good administration (mirroring articles 41-44 CFR)
Examples of AI systems currently in use	• **AVOLA** is an algorithm used for assessing eligibility for benefits in the Municipality of Rotterdam in the Netherlands. The algorithm underpinning the system is trained to review the applications received, assess if they fulfil the eligibility criteria established and advise the employees accordingly. In order to do so, AVOLA issues a report with advice on whether there is a right or not, how the legislation and regulations have been applied to the data and which data are decisive. The employee then makes the final decision. It also assists potential applicants for social

Co-funded by the European Union

	assistance benefits, so they know in advance if they are eligible or not. To do so, it is trained to receive questions from the applicant on their chances of qualifying for benefits, based on the information they provide. Therefore, the algorithm replaces a physical employee in answering questions. • The Dutch government's **Social Insurance Bank** uses a series of "algorithms" to determine if a resident is entitled or not to child benefits. The initial assessment for eligibility is done by the government without request from the resident. The relevant municipality provides data on the child and the parents, then one algorithm combines all this data, assesses it against the relevant laws and automatically determines whether it is necessary or not to send the resident an application form for child benefit. If the algorithm determines that the resident is not entitled to child benefit, the application form is not sent. Once the child benefit has been awarded, "other algorithms" are used to ensure the payment is made per quarter automatically. These algorithms also automatically process changes with the family situation, such as the birth of other children or change of residence within the country. These algorithms are prescriptive, meaning that they make decisions completely automatically, without human review and final assessment. According to the government, the use of these algorithms: better predicts how much all benefits together will cost for the National Budget; saves costs of manual processing of customer files and allows employees to focus on customisation only "where necessary"; prevent debts as it calculates the correct benefit amount. • An algorithm deployed by the **Dutch Employee Insurance Agency** (UWV), assesses the risk that an applicant for unemployment benefits is "culpably unemployed", meaning that they have become unemployed by their own fault, e.g.

Co-funded by the European Union

	by resigning from their job without a good reason. The algorithm makes the determination based on data from previous applications for benefits (e.g. how often the applicant has already applied for unemployment benefits and whether they have been previously culpably unemployed), data about the applicant's employment history and information from the current application for benefits. If the algorithm detects a high risk of culpable unemployment, the file is examined by civil servants who then make the final determination. If culpable unemployment is confirmed, the applicant is not granted unemployment benefits. • The **System Risk Indication (SyRI)** was a risk assessment tool used by the Dutch government to detect various forms of fraud, including in social benefits, allowances, and taxes. SyRI's use was discontinued in 2020, following a judgement by the Dutch District Court that found it violates the right to private and family life as enshrined in Art. 8 ECHR. • The **Estonian Unemployment Insurance Fund** uses a decision support tool, **OTT**, which predicts the likelihood of the Fund's beneficiaries to find a new job and the probability of them experiencing unemployment again, highlighting the factors affecting these probabilities. It helps counsellors get a quick overview of a client's situation and set priorities according to the client's need for assistance.[165] OTT applies the random forest machine learning model, trained and tested based on the last five years' unemployment data. It takes into account a wide range of attributes and indicators, such as the person's education, previous job experience, right to benefits, disability, as well as information relating to the labour market, such as the number and type of available positions in different regions and the number of newly

[165] https://www.kratid.ee/en/ai-use-cases

Co-funded by the European Union

	unemployed people.[166] The exact list of indicators is reviewed once a quarter by an analyst.[167] The data is collected from different public databases.[168] In addition to the risk score, the programme produces an explanation regarding which are the main indicators that the risk score is based on.[169] • The **Austrian employment agency (AMS)** uses a programme automatically attributing a score to each jobseeker based on several indicators. Depending on the score, jobseekers will land in one of three groups: group A for people who need no help in finding a new job, group B for people who might benefit from retraining, and group C for people deemed unemployable, who will receive less help from AMS and may be discharged to other institutions. The AMS algorithm has drawn criticism for violating anti-discrimination laws as it is more likely to assign an unemployed woman to a lower group even if her experience and qualifications match a man's.[170] Poland's public employment service (PSZ) deployed a similar system in 2014, which was, however, abandoned in 2018 following a ruling by the Constitutional Court. The system was criticised by both the Polish civil society and the Data Protection Authority due to lack of transparency, profiling, and discrimination against vulnerable persons.[171]
Risks of fundamental rights' violations	The risks related to the use of AI in social welfare are exemplified in the child benefits scandal in the Netherlands, reported earlier in the

[166] Nortal, "OTT – An AI-powered success story in the public sector", 5 July 2021. Available at: https://nortal.com/insights/ott-an-ai-powered-success-story-in-the-public-sector/
[167] Interview with the Centre of IT Impact studies (CITIS), 1 August 2024.
[168] Nortal, "Estonian Unemployment Insurance Fund prevents unemployment with Artificial Intelligence", 29 October 2022. Available at: https://nortal.com/insights/estonian-unemployment-insurance-fund-prevents-unemployment-with-artificial-intelligence/
[169] Interview with the Centre of IT Impact studies (CITIS), 1 August 2024.
[170] N. Kayser-Bril, Austria's employment agency rolls out discriminatory algorithm, sees no probleem, *Algorithm Watch*, 6 October 2019, https://algorithmwatch.org/en/austrias-employment-agency-ams-rolls-out-discriminatory-algorithm/
[171] J. Niklas, "Poland: Government to scrap controversial unemployment scoring system", *Algorithm Watch*. Available at: https://algorithmwatch.org/en/poland-government-to-scrap-controversial-unemployment-scoring-system/

Co-funded by the European Union

	country's chapter. The use of algorithms to detect fraud in childcare benefits applications included information about nationality as a risk factor and resulted in approximately 26.000 parents and caregivers being falsely accused of fraud, a great majority of which consisted of ethnic minorities and low-income families, whereas over 1500 children were forcefully removed from their homes. The use of AI in this area may be influenced by biases which inadvertently result in **discrimination** and **restricted access to social welfare services.** These biases often lead to indirect discrimination because of the inclusion of proxies, i.e., seemingly neutral pieces of information that are nevertheless strongly related to a protected characteristic. For example, shoe size as a proxy for gender or names as a proxy for ethnicity, as in the Netherlands case. Discrimination resulting from the use of proxies is difficult to prevent, as there is a potentially limitless number of proxies, and their correlation to a protected characteristic will be evident to various extents. [172] In addition, these algorithms often make use of **personal data**, including data related to the **private and family life** of individuals, in a disproportionate and opaque manner with serious implications for the persons concerned. This is particularly problematic where human intervention is not adequately guaranteed and where decisions are largely automated. It is worth noting that, although Article 41 CFR refers to the **right to good administration** as applicable to Union bodies, this principle is also enshrined in the common traditions of democratic states. Of particular relevance here are the obligation of the administration to give reasons for its decisions that affect individuals as well as the right of access to documents, which can be hard to ensure in some cases involving hard-to-explain algorithmic processes. At the same time, in cases where one's **right to be heard and participate in decision-making** concerning them is upheld (e.g., by requesting individuals to provide information on their claims), biases can also lead to overlooking or misinterpreting **cultural, religious and**

[172] Bias in Algorithms – Artificial Intelligence and Discrimination, European Union Agency for Fundamental Rights, 2022, ISBN 978-92-9461-935-8.

Co-funded by the European Union

	linguistic cues, disproportionately disadvantaging diverse populations.

Table 7: AI in Employment

AI in Employment	
Relevant AI capabilities	• Object detection and tracking • Emotion recognition • Collaborative robotics / human robot interaction • Anomaly / outlier detection • Logistics planning and scheduling • Forecasting • Correlation and association analysis
Charter rights potentially affected	• Article 1, Human dignity • Article 4, Prohibition of torture and inhuman or degrading treatment or punishment • Article 8, Protection of personal data • Article 15, Freedom to choose an occupation and right to engage in work • Article 21, Non-discrimination • Article 22, Cultural, religious and linguistic diversity • Article 23, Equality between men and women • Article 26, Integration of persons with disabilities • Article 27, Workers' right to information and consultation within the undertaking • Article 28, Right of collective bargaining and action • Article 29, Right of access to placement services • Article 30, Protection in the event of unjustified dismissal • Article 31, Fair and just working conditions • Article 32, Prohibition of child labour and protection of young people at work • Article 33, Family and professional life.

Co-funded by the European Union

Examples of AI systems currently in use	• **Job-matching** is increasingly performed with the use of AI systems, especially when it comes to large platforms and job search sites, such as LinkedIn, Monster, CareerBuilder, etc. The algorithms used process information from both the job seeker and the employer to curate a list of recommendations for each. **nPloy** is a mobile app that connects employers and prospective employees. [173] When an employer posts a vacancy, nPloy's built-in AI algorithm filters talents and shows the vacancy to suitable candidates only. As recruiters browse candidates' CV, applicants receive vacancies that match their criteria such as salary expectations, skills, and qualifications. The app also allows job candidates to track their job applications in real time. If an applicant's CV matches the requirements of an employer, the app provides an opportunity for chat and video calls. During the initial stages of the 'matching' process, applicants' CVs do not show personal data including name, age, gender, or picture. The data anonymization feature aims to ensure transparency and equal opportunity during the initial selection process, whereby job candidates are assessed solely on the basis of their qualifications, experience, talents, and merit. The AI algorithm does not make decisions whether a certain candidate should be invited to an interview or not but its role is limited to matching vacancies with applicants. HR officers have the final say in determining whether a candidate is a suitable match or not. **Workday** [174] is an AI-powered platform that combines HR and payroll management. Its AI-powered hiring tools were contested as discriminatory in a class action filed with the Northern District of California District Court. The Court found the company to be directly

[173] See nPloy job tool. See also nPloy, Diverse & inclusive workspace.
[174] https://www.workday.com/.

Co-funded by the European Union

	liable for employment discrimination on the basis of on the basis of race, age, and disability. [175] • **Recruitment decision-making. HireVue** [176] is one of a growing number of artificial intelligence tools that companies use to assess job applicants. The algorithm analyses video interviews, using everything from word choice to facial movements to figure out an "employability score" that is compared against that of other applicants. • **Workers' rights.** Large sharing economy platforms, such as Uber, Lyft, Deliveroo, and others, use AI to **calculate wages**, often leading to inconsistencies in pay between their workers. In the US, **Uber** has abandoned the "number of fares – distance per fare" calculating method since 2022, and replaced it with the "upfront fares" system, which is based on a standard base fare, complemented by additional "incentives", including bonuses based on "quests" (e.g., the completion of a certain number of fares within a week) and increased fares during "surges". However, these opportunities are distributed unevenly and it is unclear how the additional income they offer is calculated. For example, Uber drivers have reported that "quests are not offered every week, not everyone receives a quest when they are offered, and not everyone who is offered a quest is offered the same bonus amount." [177] This algorithmic wage discrimination becomes increasingly worrying as a growing number of workers are unable to figure how their wages are calculated. **Amazon** is using AI-powered cameras and other monitoring systems to evaluate the performance and productivity of its employees. Specifically, cameras are used to track driver behaviour and facial expressions, and are

[175] Case of Mobley v. Workday, https://www.seyfarth.com/news-insights/mobley-v-workday-court-holds-ai-service-providers-could-be-directly-liable-for-employment-discrimination-under-agent-theory.html.
[176] https://www.technologyreview.com/2019/11/07/75194/hirevue-ai-automated-hiring-discrimination-ftc-epic-bias/.
[177] For more information, see *On algorithmic wage discrimination*, Veena Dubal, Columbia Law Review, available at https://www.columbialawreview.org/content/on-algorithmic-wage-discrimination/.

Co-funded by the European Union

	being considered for tracking behavioural biometrics to enhance cybersecurity and detect data breaches. These systems have raised concerns about employee privacy and working conditions. The French Data Protection Authority (CNIL) recently fined Amazon €32 million for what it deemed excessive monitoring practices, particularly the precision with which work interruptions were measured. [178] Technology-enhanced/AI-powered workplace surveillance has become commonplace, especially among large corporations, jeopardising more rights except for the privacy of the employees, especially workers' rights such as the right to safe, healthy, and dignified working conditions, the right to collective action, and others. • **Transformation of the labour market and job loss.** Large food retailers are some of the biggest employers of low-skilled workers. Corporations like Tesco in the UK are increasingly relying on technology solutions to decrease the number of employees in their stores, including through the use of AI. Since the opening of its first check-out-free store in 2021, Tesco currently operates 4 stores in London and Birmingham where customers can shop and walkout without scanning their products, checking-out or paying on location. The stores operate without any cashiers, through an app called GetGo. [179] Generative AI has also caused significant job loss in the creative industry. A 2024 survey by the Society of Authors in the UK, [180] found that a stunning a third of translators and quarter of illustrators losing work to AI. Job loss has extended to the gaming industry, where concept artists and video game actors are protesting their gradual

[178] https://www.workyard.com/answers/what-employee-monitoring-software-does-amazon-use.

[179] https://www.tesco.com/zones/getgo?srsltid=AfmBOooD7U4NcKclTObJVvlF2dxdrct6GWyL9n4Swcjhs5borYE8SFPT.

[180] https://societyofauthors.org/2024/04/11/soa-survey-reveals-a-third-of-translators-and-quarter-of-illustrators-losing-work-to-ai/#:~:text=A%20quarter%20of%20illustrators%20(26,value%20because%20of%20generative%20AI.

Co-funded by the European Union

	replacement by AI. [181] Similar issues are plaguing different fields (e.g., the film and television workers, mentioned in the introductory section of this study).
Risks of fundamental rights' violations	Careful design of job-matching algorithms can help employers streamline their HR and match the right job candidates to their desired position -to the benefit of both. However, even the most meticulously designed systems are not foolproof against **biases**. For example, the LinkedIn algorithm was found to produce biased results and disproportionately suggest men vs. women for higher-ranking positions. [182] Despite corrective action which now excludes a person's name, age, gender, and race from the data fed into the algorithm, to avoid latent biases, the LinkedIn team found that the service's algorithms could still detect behavioral patterns exhibited by groups with particular gender identities. Recruitment tools like HireVue pose much greater risks. Conducting interviews with no human participation is bound to significantly increase AI's influence in the decision-making process of hiring one candidate vs. the other. A key issue here is the lack of a sound scientific basis behind the analysis performed by these systems. Factors like gestures, pose, lean, tone and cadence are unreliable and there is no evidence that provide credible assessments. Moreover, they are likely to reproduce biases favouring white, male, able-bodied candidates against candidates of colour, non-native speakers, persons with disabilities, etc. [183] Another issue is the lack of transparency that comes with this privately developed, proprietary tools, rendering the identification and proof of bias extremely difficult. **Workers' rights** are also profoundly affected by intrusive work-place surveillance through the use of AI. Unreliable emotion

[181] https://www.npr.org/2024/08/14/nx-s1-5072638/video-game-strike-ai-animation-sag-aftra.
[182] https://www.technologyreview.com/2021/06/23/1026825/linkedin-ai-bias-ziprecruiter-monster-artificial-intelligence/.
[183] See, e.g., *For some employment algorithms, disability discrimination by default*, Alex Engler, Brookings (2019), accessible at https://www.brookings.edu/articles/for-some-employment-algorithms-disability-discrimination-by-default/.

Co-funded by the European Union

	recognition systems, AI-powered cameras, tracking and behavioural analysis are among the many tools used to monitor and evaluate the performance of employees, in a manner that not only violates their right to **fair and just working conditions** but also their **dignity** and, at times, verges on **inhuman treatment**. Beyond their treatment, these systems can influence other workers' rights, including their right to organise and engage in collective bargaining and action, their right to information and consultation with their employer, and their right to be protected by unjustified dismissal – the latter two predominantly due to the lack of transparency of "black box" systems. Last but not least, as with all transformative technologies, AI has had a ripple effect on the job market, introducing unprecedented automation and already accounting for significant job loss. As such, **the right to engage in work** faces a clear risk which needs to be mitigated through **just transition** policies. LinkedIn recently published the top 25 fastest growing jobs - a lot have to do with AI and many, if not all, exhibit wide gender gaps. [184]

Table 8: AI in Asylum, Migration and Border Control

AI in Asylum, Migration and Border Control	
Relevant AI capabilities	• Tracking • Emotion recognition • Audio-based sentiment analysis • Drones • Forecasting • Causal inference and corelation/association analysis • Logistics
Charter rights potentially affected	• Article 1, Human dignity • Article 2 Right to life

[184] https://www.linkedin.com/pulse/linkedin-jobs-rise-2025-25-fastest-growing-us-linkedin-news-gryie/?trackingId=iBBYpYFvS5uYpaqZg%2FuxKA%3D%3D.

Co-funded by the European Union

	• Article 4, Prohibition of torture and inhuman or degrading treatment or punishment • Article 7, Respect for private and family life • Article 8, Protection of personal data • Article 10, Freedom of thought, conscience and religion • Article 11, Freedom of expression and information • Article 18, Right to asylum • Article 19, Protection in the event of removal, expulsion or extradition • Article 20, Equality before the law • Article 21, Non-discrimination • Article 22, Cultural, religious and linguistic diversity • Article 23, Equality between men and women • Article 24, The rights of the child • Article 25, The rights of the elderly • Article 26, Integration of persons with disabilities • Principle of good administration (mirroring articles 41-44 CFR) • Article 47, Right to an effective remedy and to a fair trial
Examples of AI systems currently in use	• The AI-driven systems HYPERION and CENTAUR deployed in Reception and Identification Centres and Closed Controlled Structures for asylum seekers in the island of Samos, Greece are advanced tools for asylum management and security. **HYPERION** is designed for asylum management, this system collects biometric and personal data, such as fingerprints and facial recognition scans. It links this data to individual cards used to manage access to essential services, including healthcare, food, and shelter, while tracking movements within and outside reception centres. It operates as a centralized data repository, ostensibly ensuring efficient resource distribution for asylum seekers but also facilitating movement tracking, creating a panoptic surveillance environment. **CENTAUR** is aimed at enhancing security, CENTAUR employs AI

Co-funded by the European Union

	algorithms, drones, and CCTV cameras to analyse behaviours in real time. Its purpose is to detect aggression or potential escape attempts, classifying behaviours deemed "suspicious" by its predictive algorithms. The system employs behavioural analytics alongside drone-based perimeter monitoring. Together, these systems represent a paradigm shift in the governance of migration and asylum, moving toward a surveillance-heavy model. The two systems have been a topic of contention between the Ministry of Migration and Asylum and the Hellenic DPA. In its decision of 2/4/2024, the HDPA fined the Ministry a total of 175 000 Euros on account of (a) not conducting a complete, comprehensive and coherent DPIA at the design stage of the system, thus violating 25 and 35 GDPR; (b) not been transparent and forthcoming with information about the data processing activities performed by the two systems during the HDPA's inquiry. • iBorderCtrl is an experimental EU border control system employing AI to evaluate asylum seekers' credibility through facial analysis and micro expressions. Critics argued it was inaccurate and discriminatory, risking wrongful denials. • Palantir's AI-powered analytics tools have been used by the US Immigration and Customs Enforcement for surveillance and deportation operations. These systems have drawn criticism for potentially violating the non-refoulement principle by deporting individuals to unsafe conditions.
Risks of fundamental rights' violations	Migrants, refugees and asylum seekers represent a particularly vulnerable category of persons, more often than not facing multiple vulnerabilities (due to their nationalities, religion, age, disability, sexual orientation and gender identity, and other factors). At the same time, the area of asylum and border control is increasingly politicised and treated as a matter of not only national, but also

Co-funded by the European Union

	European security. [185] This often leads to the instrumentalization of this already marginalised group, their criminalisation, and the frequent violation of their rights in light of the legal limbo they often find themselves in. Intrusive surveillance of refugee camps can deprive beneficiaries of international protection of their most basic rights to **privacy, religious and other expression, family life and other freedoms**, under fear that their applications for asylum may be on the line. The automated processing of applications or the use of dubious expression analysis technologies at any stage of the proceedings, may affect their **rights to asylum and non-refoulement**, as well their right to an **effective remedy**. In the area of border control, push-backs [186] likely aided by AI-driven monitoring and tracking systems, as well as drones, endanger migrants' **lives, integrity and dignity**, and may **expose them to a risk death penalty, torture or other inhuman or degrading treatment or punishment**.

Opportunities

AI applications in the field of **healthcare** are among the most promising when it come to advancing human wellbeing and improving access to life-saving services. Although in their nascence, several AI systems have been deployed, among others in medical research, prevention and treatment. Google's **AlphaFold** [187] can accurately predict the 3D shapes of proteins, which can facilitate drug discovery efforts. **Aidoc** [188]aims to enhance the efficiency of radiology through image analysis, assisting radiologists in detecting and prioritising abnormalities in real time, and aiding in faster, more accurate diagnoses and treatment decisions. **Butterfly iQ** [189] is a point-of-care, handheld ultrasound device, using AI to help physicians identify abnormalities fast, including in hard to reach, crowded areas, and low-resource settings, making it ideal for emergency situations or humanitarian assistance. Finally, AI has also entered the field of mental healthcare. **Wysa** is an app offering mental healthcare services, focusing on Cognitive Behavioural Therapy. The app has been evaluated in various settings, [190] and has shown

[185] https://www.theguardian.com/commentisfree/article/2024/sep/11/europe-migration-asylum-seekers.
[186] https://asylumineurope.org/reports/country/greece/asylum-procedure/access-procedure-and-registration/access-territory-and-push-backs/.
[187] https://www.theregister.com/2022/09/08/deepmind_alphafold_performance/.
[188] https://www.futurepedia.io/tool/aidoc.
[189] https://pmc.ncbi.nlm.nih.gov/articles/PMC7721766/.
[190] https://www.ncbi.nlm.nih.gov/pmc/?term=wysa.

Co-funded by the European Union

promising results, including in terms of onboarding rate, retention, and engagement. Its affordability, accessibility and anonymity represent concrete advantages that may lead to more people getting necessary help.

Public administration is another area where AI, at least some of its applications, can facilitate the exercise of rights. The digitalization of public administration is proceeding in various degrees and with great discrepancies among the Member States. AI can significantly speed up this process through linguistic capabilities such as translation, text classification, chatbots and others. It can improve access to documents and other information, contributing to increased public awareness and participation, including for persons with disabilities, linguistic minority groups, etc. **Bürokratt** serves as a chatbot network for Estonian public institutions, assisting users with routine inquiries. Future versions aim to include proactive notifications and personalized assistance. The chatbot does not currently process personal data. Should such processing take place in newer editions, appropriate safeguards should be introduced to mitigate risks to the relevant rights. **ProZorro** [191] is an open-source e-procurement system, that analyses contracting data, flags high-risk deals and irregularities, and reports them to government authorities in Ukraine, with the aim to enhance transparency and combat corruption. The system has proven to be particularly effective and is often cited as a good practice. [192] Unfortunately, there are reports of efforts to undermine it by removing certain types of procurement from its scope of application (e.g., procurement for reconstruction projects after the war). [193] Many member states face challenges in terms of the speedy and effective administration of justice. Despite justice being a high-risk area, there are certain applications of AI that are encouraged, in order to promote the effective exercise of the relevant rights. Specifically, CEPEJ [194]encourages the use of AI machine learning techniques and natural language processing for keyword or full-text search of case law; chatbots that could enhance access to legal knowledge for the general public, as well as document templates (e.g., court applications, lease agreements, etc.); and the use of AI to draw up strategic approaches that can help improve the efficiency of justice, e.g., by carrying out quantitative and qualitative evaluations, making projections, and proposing key performance indicators.

Despite these systems being imperfect, they do showcase the potential of AI to drive progress, respect for fundamental rights and democratic institutions, and to improve equity across the globe, provided that is designed, developed, deployed and used responsibly, in line with strict, rights-based

[191] https://ti-ukraine.org/en/project/public-procurement-oversight/.
[192] https://www.hks.harvard.edu/publications/overcoming-corruption-and-war-lessons-ukraines-prozorro-procurement-system.
[193] https://ti-ukraine.org/en/blogs/rebuilding-with-out-prozorro-how-to-conduct-procurement-so-that-there-are-no-questions/.
[194] European ethical Charter on the use of Artificial Intelligence in judicial systems and their environment, CEPEJ (2018), accessible at https://www.coe.int/en/web/cepej/cepej-european-ethical-charter-on-the-use-of-artificial-intelligence-ai-in-judicial-systems-and-their-environment.

Co-funded by the European Union

regulatory approaches and with due respect to all ethical considerations applicable generally or in specific sectors. To this end, appropriate mitigating measures should be in place.

Mitigating measures

With the AI transformation already on the way, and both the public and – crucially – the private sector investing vast sums in the development and deployment of AI systems in all sectors, measures must be introduced in law, policy, and practice, so that the above-described risks are mitigated. Below follows an indicative list of such measures, which should be combined to ensure comprehensive protection.

Legal and Regulatory Measures

1. **AI-Specific Legislation.** Although the AI Act Regulation has direct effect for the EU Member States, its scope of application, excluding, for example national security or introducing exceptions from various obligations for, e.g., law enforcement and criminal investigations, migration, border control or asylum, renders the introduction of national laws that would fill in these gaps essential for the protection of fundamental rights. In this respect, while the AI Act adopts a risk-based approach, Member States should be encouraged to opt for a rights-based approach that would better align with the goal for human-centric AI. This could and should include laws that safeguard the workforce from undue displacement.
2. **Consistent and robust enforcement of other relevant legislation.** It has been clearly shown that legislation which may not be directly or explicitly linked to AI, such as data protection legislation or consumer protection, can be very useful in addressing many of the above-identified risks related to the deployment and use of AI systems. Anti-discrimination laws may also need to be reviewed to address new challenges arising from the use of AI.
3. **Fundamental Rights Impact Assessments (FRIAs).** States should be encouraged to introduce comprehensive frameworks for FRIAs, defining the specific procedural requirements for their implementation, in line with Article 27 of the AI Act and international ethical standards. These frameworks should guarantee human oversight in decision-making, as well as wherever else it is required for the responsible operation of the AI system.
4. **Promote the application of principles for ethical AI in the public sector.** When it comes to public administration, States should be encouraged to go further than the obligations introduced by the AI Act mandate. For example, to promote transparency, traceability, explainability and contestability of the AI systems deployed in the public sector, they should opt for “white box” AI systems whenever possible. Ethical and transparent procurement processes

Co-funded by the European Union

are also crucial to ensure that AI bought from private companies does not raise any significant ethical concerns.

Governance and Institutional Measures

7. **Independent AI Oversight Bodies**. The independence of the administrative bodies designated with AI oversight should be guaranteed, among others by ensuring the personal and functional independence of their members, the designation of adequate funds and human resources for their mission, and the prohibition of political or other interference with their work.
8. **Ethical AI Certification**. Any Ethical AI certification schemes introduced must conform with the requirements of Article 29 of the AI Act.
9. **Redress Mechanisms.** States should ensure that citizens and residents have effective remedies at their disposal to challenge AI-driven/assisted decisions, including at the administrative level. The right to human intervention should be guaranteed.
10. **Investment in infrastructure and AI readiness.** Promoting AI for good requires that the people are actually able to benefit from AI-powered solutions and that they have equal access to their use. This is often not the case, both among different states as well as within them. Unequal access to resources, such as the internet or electricity, risks leaving behind entire communities. States should address these gaps and ensure AI readiness across their populations.

Technical and Design Measures

11. **Respect for the principles of data protection.** States must ensure that the GDPR is adhered to in every case an AI system may process personal data, as well as that such processing is not performed unless strictly necessary. This includes undertaking technical and organizational measures to ensure data security by design and by default, as well as support for GDPR enforcement performed by DPAs and courts, whether concerning the public or the private sector.
12. **Bias Mitigation.** The mitigation of biases involves both the human factor (AI developers), as well as the design and programming of the AI software in ways that are conducive to it identifying and ignoring such biases when present in their training data. States should make sure that laws containing rules for ethical AI are applied rigorously. These should include laws on data protection, consumer protection, transparency, anti-discrimination, etc. In addition to these legal solutions, states should also consider issuing appropriate, practical guidance for developers, debiasing standards, and bias benchmarks.

Co-funded by
the European Union

13. **Establishment of human-centric AI sandboxes.** AI sandboxes must be established in all Member States by August 2026, in accordance with Article 57 of the AI Act. These are meant to both encourage innovation by providing a safe space for the development of AI systems with minimal risks but should also function as a pilot for these systems in terms of the potential risks they pose to fundamental rights and the welfare of people, as well as the mitigating measures needed to ensure their safe and responsible operation.

Support to the general public and the civic society

15. **AI Literacy and Public Awareness**. Public outreach and awareness to inform the public on the functions, opportunities and risks posed by AI is a key step in ensuring "no one is left behind" when it comes to the AI transformation. Moreover, informed rights bearers are more likely to pursue their rights and push toward improve protection through litigation and collective action.
16. **Training**. To further promote a just transition in the labour market, continuous training and lifelong learning for digital and AI literacy, including for those in unemployment, must be introduced. In addition, those active in relevant fields (e.g., AI developers, deployers, etc.) should be encouraged to undergo ethical and fundamental rights training and sensitization.
17. **Support for Civil Society and Watchdog Groups.** An open and free civic space is key to promote awareness and shed light to potential violations of the legal and regulatory framework. Civil Society and Watchdog Groups that are well-supported and have a good working relationship with the authorities can act as a major mitigating factor for AI-related risks. CSOs should be empowered and supported to participate in decision-making regarding AI at all levels, including national, regional and local. States are encouraged to acknowledge their role in providing valuable insights on neglected policy areas and representing vulnerable groups. Their work should be facilitated through funding, visibility of their contributions and an enabling legal and regulatory environment. [195]

[195] To this effect, see *AI Governance: Empowering Civil Society*, Renaissance Numérique (2025), accessible at https://www.renaissancenumerique.org/en/publications/ai-governance-empowering-civil-society/.

Co-funded by the European Union

Conclusions

This study has sought to highlight current and potential risks to fundamental rights as enshrined in the Charter of Fundamental Rights of the European Union. Our aim was not to underestimate the potential of AI for good, but to provide a counterpoint to the often-dominant focus on its potential for innovation and financial growth, without adequate consideration of the risks it may entail. As global inequalities deepen and fundamental rights face serious backsliding, it is important to remember that growth must benefit all in order to be equitable.

In this respect, the practice of AI to date raises serious concerns. AI used for profit, if left unchecked, is likely to lead to fundamental rights and ethical concerns being disregarded or relegated to the background. In this area, certain applications of AI are inherently worrying. Deep fakes and microtargeting used to manipulate voters or consumers are in themselves egregious, challenging legal and ethical standards related to freedom of conscience, expression and information, privacy and data protection, consumer rights, and many others. Generative AI, while promising, needs to be closely monitored due to its large-scale impact and potential incorporation of existing societal biases.

It is telling, however, that some of the most serious rights violations reported above are at the core of state activity. The increased use of AI in policing and criminal investigations, as well as in border control, reflects the political shift towards securitisation, with public safety taking precedence over respect for fundamental freedoms. In particular, the exclusion of AI used for vaguely defined national security purposes from the scope of the AI Act, and the broad exceptions to the permissible use of 'real-time' remote biometric identification systems, are indicative of this. Another area of concern is social security and welfare. So far, states appear to be more concerned with identifying fraud and cutting costs than with helping beneficiaries, leading to gross violations of fundamental rights and discrimination due to algorithmic bias.

Against this backdrop, the risk-based approach fostered by the AI Act appears to be a step backwards from previous, bolder regulations relevant to new technologies, such as the GDPR. Political considerations and balances aside, the failure to adopt a rights-based approach in this instance could be seen as a missed opportunity.

Nevertheless, promising practices point the way to harnessing the potential of AI for good. Medical applications can reduce inequalities in access to healthcare for the most vulnerable by reducing cost and geographical barriers. Similarly, applications such as ProZorro can work for the benefit of many by targeting corruption and improving transparency as a key component of well-functioning democracies. That said, the vast majority of AI applications will not fall into the "good versus evil" dichotomy. AI is a tool that can be used to improve our daily lives and collective prosperity,

Co-funded by
the European Union

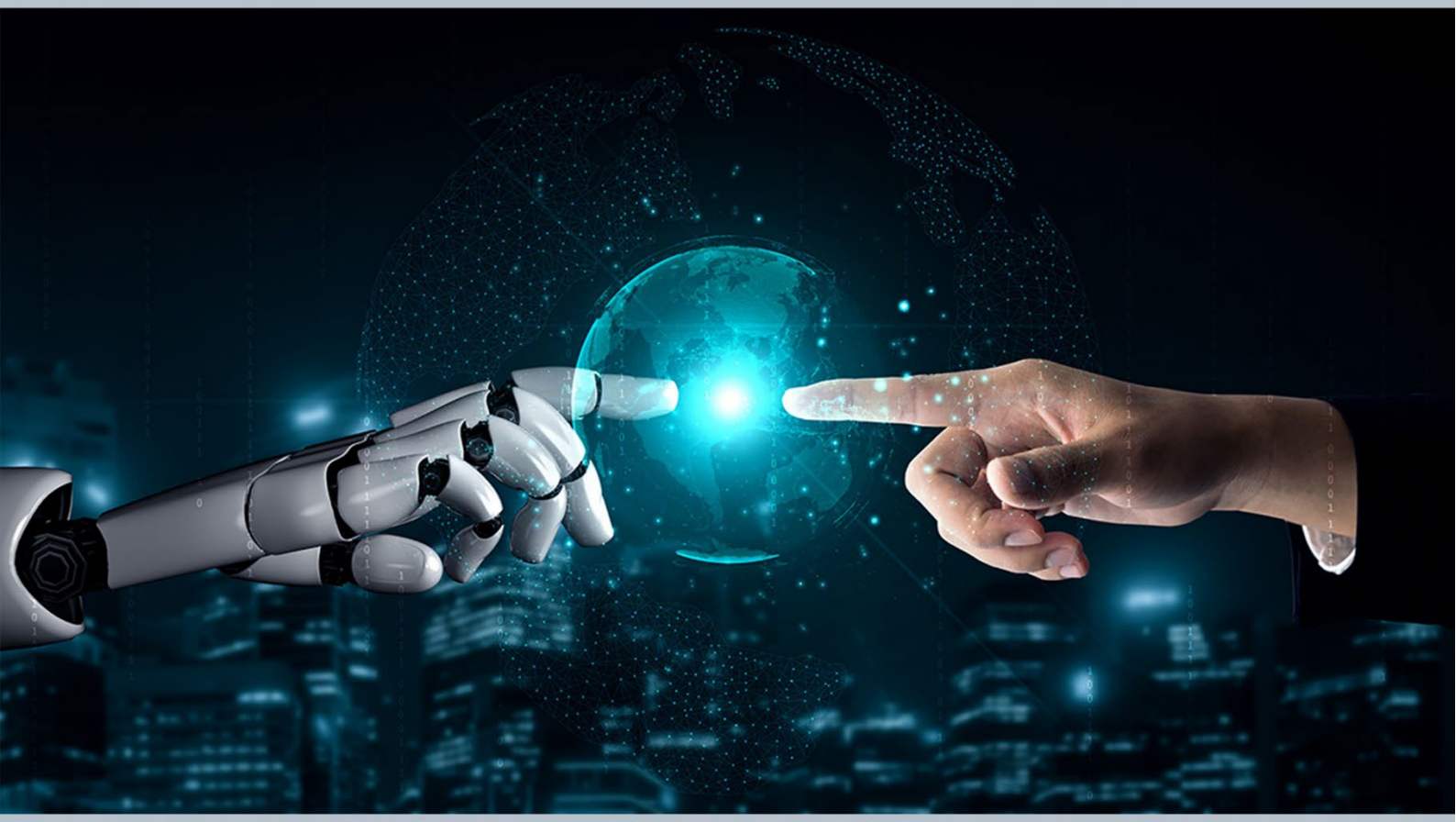
provided that its human-centred nature is prioritised at all times. To this end, experience has shown that a robust regulatory framework is necessary and that concerns about potential stifling of innovation are exaggerated.

Instead, risk mitigation measures need to be applied to prevent fundamental rights violations. To be effective, these measures must be multi-level and involve a wide range of stakeholders, including the state, social partners, academia, civil society and the general public. Education and awareness-raising are key for empowered rights holders, while technical measures are essential to prevent violations. The mandated FRIAs and the pilot testing of high-risk systems in AI sandboxes are a positive component of the AI Act that can yield results if complemented by robust monitoring and enforcement.

In conclusion, ethical, responsible and human-centred AI requires an active and engaged public, a vibrant civil society, and willing public and private actors to ensure that this ground-breaking transition leaves no one behind.

Co-funded by
the European Union

refrAIme



Co-funded by the European Union

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or EACEA. Neither the European Union nor the granting authority can be held responsible for them.