Project title	**Achilles Human-Centred Machine learning: lighter, clearer, safer**		
Project acronym	ACHILLES		
GA number	101189689		
Project start date	01/11/2024	Duration	48 months

D4.1 – LEGAL AND ETHICAL MAPPING

Due date	30/04/2025	Delivery date	30/04/2025
Work package	WP4		
Responsible Author(s)	Victoria Hendrickx (KUL)		
Contributor(s)	Viltė Kristina Dessers (KUL)		
Reviewer(s)	Gemma Galdón Clavell, Mariana Carvajal Sojo (ETICAS), André Carreiro (FhAICOS)		
Version	1.0		
Dissemination level	Public		

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

VERSION AND AMENDMENTS HISTORY

Version	Date (DD/MM/YYYY)	Created /Amended by	Changes
1.0	03/04/2025	Victoria Hendrickx	Initial draft
1.1	29/04/2025	Victoria Hendrickx	Reviewed based on feedback

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

TABLE OF CONTENTS

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ALTAI	Assessment List for Trustworthy Artificial Intelligence
CEDS	Common European Data Spaces
CEN-CENELEC	European Committee for Standardization – European Committee for Electrotechnical Standardization
CFR	Charter of Fundamental Rights of the European Union
CJEU	Court of Justice of the EU
D	Deliverable
DA	Data Act
DGA	Data Governance Act
DMA	Digital Markets Act

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

DMP	Data Management Plan
DPIA	Data Protection Impact Assessment
DPO	Data Protection Officer
DSA	Digital Services Act
ECHR	European Convention on Human Rights
ECtHR	European Court of Human Rights
EHDS	European Health Data Space
EU	European Union
FRA	Fundamental Rights Agency
FRAND	Fair, reasonable, and non-discriminatory
FRIA	Fundamental Rights Impact Assessment
GDPR	General Data Protection Regulation
GPAI	General purpose AI
HERA	Holistic Evaluation and Regulatory Adherence
HIIL	Human-in-the-loop
HLEG AI	High-Level Expert Group on Artificial Intelligence
IP	Intellectual property
LLM	Large language model
MDR	Medical Devices Regulation
ML	Machine learning
OECD	Organisation for Economic Co-operation and Development
ICCPR	International Covenant on Civil and Political Rights
ICESCR	International Covenant on Economic, Social and Cultural Rights
IDE	Integrated Development Environment

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

IEC	International Electrotechnical Commission
ISO	International Organisation for Standardization
DMA	Digital Markets Act
DSA	Digital Services Act
R&D	Research and development
SCRIPTA	Scripts Certification and Review. Integrity of Processes and Trust Assurance
VLOP	Very large online platform
VLOSE	Very large online search engine
UNESCO	United Nations Educational, Scientific and Cultural Organization
XAI	Explainable AI

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

1 EXECUTIVE SUMMARY

Deliverable D4.1 presents the initial legal analysis of legal and ethical frameworks relevant to the ACHILLES project.

The Deliverable focuses on the legal requirements for AI development at both the EU and international levels, with particular attention on fundamental rights and privacy and data protection. It also examines broader European legislative instruments concerning data governance, information society services, and cybersecurity while further looking into some sector-specific legal requirements. It ends with a discussion on ethical considerations.

Chapter 2 introduces the ACHILLES project, its objectives and methodologies, and outlines the purpose of this deliverable, i.e. establishing the initial compliance roadmap by identifying and interpreting relevant legal frameworks.

Chapter 3 focuses on the importance of fundamental rights within the ACHILLES project. It highlights the most relevant rights for the project and emphasises the need to respect these rights throughout all activities. This chapter also explores how automation and AI may necessitate a broader or more specific interpretation of these rights to address the emerging risks that the technologies bring with them, which may not have existed in traditional contexts. It subsequently focuses on the right to human dignity and integrity of the person, respect for private and family life and the protection of personal data, equality and non-discrimination, intellectual property, healthcare, and environmental protection.

Chapter 4 outlines AI-related requirements, with a focus on the European AI Act, given its direct and extensive impact. The chapter provides an overview of the AI Act's classification rules and obligations most relevant to the project. It also emphasises the importance of international initiatives and standards.

Chapter 5 addresses privacy and data protection requirements by examining the General Data Protection Regulation in detail. It reiterates the core principles and rights and considers how ACHILLES activities can align with these rules. It also discusses the processing of special categories of personal data, such as health-related and biometric data, the status of synthetic data, the appointment of a data protection officer, and data protection impact assessments.

Chapter 6 presents an overview of relevant data governance regulations, i.e. the Data Act, Data Governance Act and Common European Data Spaces. They aim to facilitate reliable and secure access to data. Given that data is a central element in various ACHILLES activities, such as the development of AI systems, these frameworks are important as they provide a regulatory framework to help manage data access and sharing in a compliant manner.

Chapter 7 focuses on the regulation of information society services under the Digital Services Act and the Digital Markets Act. Their shared objective is to create a secure and transparent digital environment

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

that protects users' fundamental rights and ensure a level playing field for businesses, but also promote innovation, growth and competitiveness.

Chapter 8 discusses cybersecurity regulations that are important for ensuring the secure development of AI systems and models in the project.

Chapter 9 explores the relevance of sectoral legislation and identifies some laws that could be triggered by ACHILLES activities. It emphasises the need for further research in the coming months to refine the approach to sectoral legislation and determine the appropriate level of compliance checks.

Chapter 10 provides an overview of some ethical considerations for the project, including the need to adhere to trustworthy AI principles, address the risks associated with generative AI and ensure the ethically compliant involvement of participants in the use case validation.

Finally, Chapter 11 concludes the deliverable by summarising the relevant frameworks identified for the project and their requirements.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

2 INTRODUCTION

ACHILLES starts from the observation that trustworthiness and efficiency are two of the most significant Achilles' heels of machine learning (ML). The project aims to develop a **modular framework** that fosters the creation of AI-based systems that are lighter, clearer and safer. By drawing inspiration from the iterative processes of clinical trials, ACHILLES seeks to establish an AI development ecosystem that is efficient, compliant and trustworthy.

The main objective of ACHILLES is to enable the development of energy- and data-efficient ML models that maintain high performance while ensuring trustworthiness, positive societal impact, and reduced environmental footprint. To achieve this, the project will deliver the ACHILLES **Integrated Development Environment** (IDE), a user-friendly platform designed to support both beginner and expert developers through all stages of ML development.

The ACHILLES IDE will serve as a comprehensive framework, embedding the project's findings into AI solutions that prioritise transparency, efficiency and ethical integrity. It will provide an environment to design and implement robust, compliant and human-centred AI systems through iterative development cycles. Central to the user experience is a virtual assistant (copilot), an LLM-based assistant integrated into the user interface to streamline the development process by offering real-time support, including project navigation assistance, best practice suggestions, and easy access to relevant technical, ethical, and legal documentation. It ensures that developers have a conversational partner throughout the AI development lifecycle, from co-design to monitoring after deployment. The IDE will support plugin integration, making it more adaptable to various domains and use cases.

The ACHILLES IDE consists of several interconnected modules, each supporting different stages of AI development. Figure 1 provides a visual representation of the ACHILLES IDE's architecture.

Figure 1. ACHILLES Integrated Development Environment (IDE) architecture overview (source: proposal)

To ensure the ACHILLES IDE meets practical compliance and trustworthiness standards, **use cases** will be defined and tested in collaboration with end-users and stakeholders from different domains. The use cases serve to validate the IDE's functionality and must demonstrate its ability to provide relevant compliance insights while also guiding the platform's further refinement.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

The first use case concerns **SCRIPTA** (‘*Scripts Certification and Review. Integrity of Processes and Trust Assurance*’), which focuses on AI-generated script validation, ensuring compliance with editorial, ethical, and legal standards. The use case aims to certify content integrity and prevent misuse of AI in literacy and media production. The second use case is situated in the context of **healthcare diagnostics** and examines the development and deployment of ML-based screening and diagnostic tools for ophthalmological disease detection from eye fundus scans and relevant clinical data. The third use case, **HERA** (‘*Holistic Evaluation and Regulatory Adherence*’), is designed to support efficiency and regulatory compliance within the pharmaceutical sector. The fourth use case concerns **automated identity verification** and focuses on enhancing the reliability and privacy compliance of deep learning models used in identity document verification.

Deliverable 4.1 presents the initial legal analysis of legal and ethical frameworks relevant to the ACHILLES project. The corresponding Task 4.1 is led by KUL and the contributors are ETICAS, AXIOLOGIC, CTTI and CUOMOIT. D4.1 establishes the initial compliance roadmap by identifying and interpreting international and EU legal frameworks relevant to the project activities. The full delineation of legal requirements, in-depth “why”’s and possible measures will be further examined in an updated report in D4.2 (M48). Nevertheless, D4.1 already raises some emerging research questions that warrant further examination in the future.

The Deliverable focuses on requirements for AI development at both EU and international level, with a particular focus on fundamental rights and privacy and data protection. It also examines broader European legislative instruments concerning data governance, information society services, and cybersecurity, while further looking into some sector-specific legal requirements. It ends with a discussion on the ethical considerations for the project. For a more comprehensive ethical analysis, reference should be made to D4.4 “*Ethics Guidelines (v1)*”, led by ETICAS. While D4.1 and D4.4 are closely linked and complementary, the latter focuses specifically on ethical guidelines for responsible research within ACHILLES.

Funded by the European Union
This project has received funding from the European Union’s Horizon Europe research and innovation programme under Grant Agreement No 101189689.

3 FUNDAMENTAL RIGHTS

3.1 The importance of fundamental rights in ACHILLES

Respect for fundamental rights is key for all activities within ACHILLES – both for the ACHILLES IDE itself and the AI systems developed through the IDE, not only because respecting fundamental rights is a legal obligation but also because it is a matter of ethical and social responsibility, and fundamental rights embody important normative principles that uphold democratic values. Ensuring respect for fundamental rights mitigates potential risks – whether legal, ethical, or social – and helps prevent unintended harms. It also fosters trust among individuals and society at large, enhances the credibility of project outcomes and contributes to the overall fairness and integrity of project activities.[1]

This chapter provides an overview of fundamental rights most relevant to ACHILLES and that require special attention throughout the project. It discusses fundamental rights' legal foundations and their importance – especially in the context of automation – and in light of the ACHILLES activities. The chapter begins by introducing the main legal sources of fundamental rights.

3.2 Brief introduction to the legal sources of fundamental rights

3.2.1 European Convention on Human Rights

The European Convention on Human Rights (ECHR)[2] is one of the main legal sources of human rights protection in Europe. It is part of the Council of Europe, an international organisation comprising 46 Member States – including all 27 EU countries as well as non-EU states. It was adopted in 1950 and entered into force in 1953.[3]

The ECHR focuses on safeguarding human rights, democracy, and the rule of law across Europe. While it mainly focuses on civil and political rights, it also includes certain economic and social rights. The European Court of Human Rights (ECtHR) is responsible for overseeing the implementation of the ECHR. If individuals believe their rights under the ECHR have been violated, they can bring individual complaints before the ECtHR, if all domestic legal remedies are exhausted.

3.2.2 Charter of Fundamental Rights of the European Union

[1] Karp, D.J., 2020, 'What is the responsibility to respect human rights? Reconsidering the 'respect, protect, and fulfill' framework', *International Treaty*. CUP, 12(1), pp. 83-108

[2] Council of Europe, 1950, *European Convention on Human Rights*. https://www.echr.coe.int/documents/d/echr/Convention_ENG.

[3] Council of Europe, 2025, A Convention to protect your rights and liberties. https://www.coe.int/en/web/human-rights-convention/home

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

At the EU level, the Charter of Fundamental Rights (CFR)[4] serves as a key legal instrument for the protection of fundamental rights. It applies to citizens and residents of EU Member States and entered into force in 2009.

The CFR focuses on a broad range of rights, drawing from constitutional traditions and international obligations common to EU Member States. Compared to the ECHR, the CFR has a slightly wider scope and encompasses civil, political and economic rights. The CFR is legally binding on EU institutions, bodies and agencies in all their actions, as well as on Member States when they are implementing EU law. This ensures that individuals and legal entities are protected against fundamental rights violations by EU institutions. In case of violation, the Court of Justice of the EU (CJEU) has the authority to review the legality of legislation.

As for the relation between the CFR and ECHR, all EU Member States are bound by the ECHR. In fact, for many years, the EU has expressed its intention to accede to the ECHR, so as there would be a direct link between the two systems – currently, only the individual EU Member States are party to the ECHR. Despite this, there is already a strong connection between the two frameworks. Art. 53 CFR explicitly states that nothing in it should be interpreted as restricting or adversely impacting human rights recognised in the ECHR and other human rights treaties. Moreover, the CJEU often refers to the ECHR and the case law of the ECtHR when interpreting fundamental rights within EU law. Rather than competing, the ECHR and CFR are complementary legal frameworks, strengthening human rights protection in Europe. They coexist and provide different avenues for the enforcement of fundamental rights and ensure comprehensive legal protection at both the EU and broader European levels.[5]

3.2.3 International Covenant on Civil and Political Rights

The International Covenant on Civil and Political Rights (ICCPR) is a legally binding international treaty focusing on protecting civil and political rights of individuals.[6] It was adopted by the United National (UN) General Assembly in 1966 and entered into force in 1976. The UN Human Rights Committee monitors the implementation and compliance with the provision of the ICCPR.[7]

The ICCPR guarantees a wide range of fundamental rights and freedoms, such as the right to self-determination, equality, life, liberty, security, fair trial, privacy, freedom of expression, association,

[4] Charter of Fundamental Rights of the European Union, OJ C 326, 26.10.2012, p. 391–407.
[5] Brittain, S., 2015, 'The Relationship Between the EU Charter of Fundamental Rights and the European Convention on Human Rights: an Originalist Analysis', *European Constitutional Law Review*, 11(3), pp. 482-511; Douglas-Scott, S., 2015, 'The Relationship between the EU and the ECHR Five Years on from the Treaty of Lisbon', *Five Years Legally Binding Charter of Fundamental Rights*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2533207.
[6] OHCHR, 1966, *International Covenant on Civil and Political Rights*. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights.
[7] OHCHR, 2025, *Human Rights Committee*. https://www.ohchr.org/en/treaty-bodies/ccpr; OHCHR, *Civil and Political Rights: The Human Rights Committee, Fact Sheet No. 15*. https://www.ohchr.org/sites/default/files/Documents/Publications/FactSheet15rev.1en.pdf.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

assembly and movement. In addition to the main Covenant, two Optional Protocols supplement its provisions. The First Optional Protocol establishes a mechanism that allows individuals to submit complaints to the UN Human Rights Committee regarding violations of the ICCPR.[8] The Second Optional Protocol aims to abolish the death penalty.[9] Both Protocols are optional for countries to adopt.

3.2.4 International Covenant on Economic, Social and Cultural Rights

The International Covenant on Economic, Social and Cultural Rights (ICESCR) is also a legally binding international treaty, but it specifically focuses on protecting the economic, social and cultural rights of individuals.[10] Like the ICCPR, the ICESCR was adopted by the United National (UN) General Assembly in 1966 and entered into force in 1976. The UN Committee on Economic, Social and Cultural Rights monitors the implementation of and compliance with the ICESCR.[11]

The ICESCR protects a wide range of rights and freedoms, such as the right to non-discrimination, equality, work, formation of trade unions, strike, social security, water and sanitation, protection from economic exploitation, health, education, cultural life, benefits of science, undertake scientific research and creative activity. State Parties can opt to become party to the Optional Protocol to the ICESCR, which establishes an individual complaint mechanism similar to the First Optional Protocol to the ICCPR, allowing individuals to submit complaints regarding violations of the provisions of the ICESCR.

3.3 Human dignity and the right to the integrity of the person

Human dignity and the right to the integrity of the person, as enshrined in Art. 1 and 3 CFR and Art. 10 ICCPR[12], are the first set of fundamental rights that must be respected throughout ACHILLES. Given their connectedness, they are discussed together.

There is no universally accepted definition of human dignity. However, it is often associated with principles such as self-respect, the intrinsic value of individuals, and their worth as human beings. Crucially, human dignity is inherent and inalienable, belonging to every individual by virtue of being

[8] OHCHR, 1966, *Optional Protocol to the International Covenant on Civil and Political Rights*, https://www.ohchr.org/en/instruments-mechanisms/instruments/optional-protocol-international-covenant-civil-and-political.
[9] OHCHR, 1986, *Second Optional Protocol to the International Covenant on Civil and Political Rights, aiming at the abolition of the death penalty*. https://www.ohchr.org/en/instruments-mechanisms/instruments/second-optional-protocol-international-covenant-civil-and.
[10] OHCHR, 1966, *International Covenant on Economic, Social and Cultural Rights*. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights.
[11] OHCHR, 2025, *Background to the Covenant*, https://www.ohchr.org/en/treaty-bodies/cescr/background-covenant; OHCHR, 2025, *Committee on Economic, Social and Cultural Rights*, https://www.ohchr.org/en/treaty-bodies/cescr.
[12] Art. 10 ICCPR is specifically concerned with the right to dignity in situations where individuals are deprived of their liberty.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

human and cannot be lost, forfeited or taken away, even when it is violated. It includes protection from humiliation, insults and degrading treatment and implies that individuals must always be treated with fairness and respect. Human dignity underpins the recognition of fundamental moral and political rights and duties.[13] It is an overarching principle that must be upheld, even in situations where other fundamental rights are restricted.[14]

Closely linked to human dignity is the right to integrity of the physical and mental person. While there is no universally accepted definition of what constitutes integrity, it can include the prohibition of torture, ill-treatment, inhuman or degrading treatment, violation of personhood or loss of control and belief in one's own self-efficacy.[15] In the field of medicine, Art. 3 CFR specifies that this right implies that free and informed consent is required for medical procedures and that certain unacceptable medical practices are prohibited, namely reproductive cloning of human beings, making human bodies or its parts a source of financial gain and eugenic practices.

In the context of ACHILLES, it is essential to ensure the continued protection of human dignity and the integrity of the person throughout all project activities, especially when AI is involved.

Studies have highlighted that AI-driven systems may necessitate additional or more specific safeguards beyond those traditionally in place.[16] For instance, the EU Fundamental Rights Agency (FRA) has emphasised that AI-driven processing of personal data must be carried out in a manner that respects human dignity and that this implies a **human-centred approach** where individuals remain at the centre of all AI-related discussions. Human dignity and the integrity of the person also dictate that individuals must not be subjected to AI systems without their knowledge or informed consent. Human dignity and integrity may entail specific ramifications in certain contexts or sectors. For instance, in healthcare, it requires that the use of AI should not depersonalise medical treatment – which might be relevant for healthcare use cases. Similarly, in scientific research, in particular when human participants are involved, the protection of human dignity and the integrity of the person are crucial[17]– the latter further explained in D4.4.

[13] Debes, R., 2023, 'Dignity', *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/dignity/; Tasioulas, J., 2015, *Human Dignity and the Foundations of Human Rights*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2557649#:~:text=This%20chapter%20investigates%20whether%20human%20rights%20are%20grounded,human%20dignity%3A%20the%20deontological%20and%20the%20personhood%20objections.
[14] FRA, 2025, *Article 1 – Human Dignity*, https://fra.europa.eu/en/eu-charter/article/1-human-dignity.
[15] Tiedemann, P., 2023, 'Human Rights Concerning the Protection of Physical and Mental Integrity', *Philosophical Foundation of Human Rights*. Springer, pp. 141-158.
[16] FRA, 2020, *Getting the future right – Artificial intelligence and fundamental rights,* https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf.
[17] NHB, 2022, 'Science must respect the dignity and rights of all humans'. *Natural Human Behaviour* 6, pp. 1029–1031.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

3.4 Respect for private and family life and the protection of personal data

The right to respect for private and family life, and the protection of personal data, as enshrined in Art. 8 ECHR and Art. 7-8 CFR, guarantees that every individual has the right to privacy in their private and family life, home, correspondence and communications, as well as the right to the protection of their personal data. Given their interconnectedness, these rights are examined together. They both aim to safeguard similar values, namely autonomy and human dignity, by ensuring a personal sphere in which individuals can develop their personalities, express their thoughts and form opinions freely. As such, these rights serve as a prerequisite for the exercise of other fundamental rights, such as freedom of expression, assembly and religion.[18]

The ECtHR has consistently interpreted the notions of 'private life' and 'privacy' broadly, including both physical and psychological integrity, as well as a person's physical and social identity. It also entails a range of personal attributes, such as one's name, image and the right to know one's origin. Privacy can sometimes extend to the public sphere, covering professional, commercial and financial interests, and the right to establish and maintain relationships with others.[19] It entails protection from arbitrary interference by both public authorities and private entities. However, it is not an absolute right and may be subject to restrictions.

The right to personal data protection ensures that individuals' data is processed fairly, for specific and legitimate purposes, and based on consent or another lawful basis. It grants individuals the right to access their personal data and the right to rectify it when necessary. These principles are further detailed in the General Data Protection Regulation (GDPR), which provides a comprehensive framework for data protection in the EU (*infra* 5.1).

The widespread use of AI technologies has introduced new challenges related to the collection, storage and processing of personal data. In the digital age, vast amounts of personal data are increasingly collected and processed in complex and opaque ways, which pose significant risks to privacy and data

[18] FRA, 2018, *Handbook on European data protection law*, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG.
[19] FRA, 2020, *Getting the future right – Artificial intelligence and fundamental rights,* https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf; ECtHR No. 42326/98, Odièvre v. France, 13 February 2003; ECtHR No. 77924/01, Albanese v. Italy, 23 March 2006; ECtHR No. 13444/04, Glor v. Switzerland, 30 April 2009.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

protection rights.[20] Certain categories of personal data, such as health-related or biometric data, require heightened protection due to their sensitive nature.[21]

3.5 Equality and non-discrimination

Art.14 ECHR and Protocol No. 12 to the ECHR[22], Art. 20-21 CFR, Art. 26 ICCPR, Art. 2 ICESCR and Art. 18 TFEU protect the fundamental right of equality and non-discrimination. Non-discrimination ensures that individuals have equal and fair opportunities available in society.[23]

Discrimination is often classified in two ways: **direct and indirect discrimination**. Direct discrimination occurs when individuals who are in a similar situation and should receive similar treatment are treated less favourably simply because of a particular protected characteristic or ground. Protected grounds are identifiable, objective or personal characteristics, or status by which individuals or groups are distinguished from each other. Art. 21 CFR explicitly lists several protective grounds on which cannot be discriminated, namely sex, race, colour, ethnic or social origin, genetic features, language, religion or belief, political or other opinion, membership of a national minority, property, birth, disability, age or sexual orientation. [24] In contrast, Art. 14 ECHR adopts a more open-ended list of protective grounds developed on a case-by-case basis and states that there shall be no discrimination "*on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status*." The ECtHR has used the category of 'other status' to expand protection to other grounds such as disability, age and sexual orientation. Indirect discrimination arises when a seemingly neutral rule disadvantages a person or a group of persons as a result of their particular characteristics. Indeed, biases may still emerge through proxy variables, i.e. data points that correlate with protected characteristics, such as postcode or language, which can serve as indirect indicators of race, socio-economic status or ethnicity.

Discrimination can become more pervasive in the context of automation and AI due to algorithmic biases that reinforce or exacerbate existing inequalities. Discriminatory outcomes in AI systems can stem from both data bias and model bias, often leading to indirect discrimination. Several mechanisms

---

[20] FRA, 2020, *Getting the future right – Artificial intelligence and fundamental rights,* https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf; FRA, 2018, *Handbook on European data protection law*, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG; ECtHR No. 30562/04, Marper v. UK, 4 December 2008.

[21] ECtHR No. 7508/02, L.L. v. Latvia, 10 October 2006; ECtHR No. 52019/07, L.H. v. Latvia, 29 April 2014; ECtHR, 2024, *Factsheet – Personal data protection*, https://www.echr.coe.int/documents/d/echr/FS_Data_ENG.

[22] Protocol No. 12 to the ECHR introduces the prohibition of discrimination as a free-standing right, as opposed to Art. 14 ECHR that prohibits discrimination only in relation to the exercise of another right in the ECHR.

[23] ECtHR No. 31871/96, Sommerfeld v. Germany, 8 July 2003; ECtHR No. 17209/02, Zarb Adami v. Malta, 20 September 2006; ECtHR No. 6033/13, A.H. v. Russia, 17 January 2017; FRA, 2018, *Handbook on European data protection law*, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG.

[24] FRA, 2020, *Getting the future right – Artificial intelligence and fundamental rights,* https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

can contribute to this issue.[25] In the first place, AI models may rely on proxy variables that unintentionally encode biases. For example, while postcode is not a protected characteristic, it may be strongly correlated to race, leading to indirect discrimination in decision-making such as credit scoring or law enforcement. Second, biases can result from biased training data. If the training data of AI models reflect societal biases, lack diversity, or disproportionately represent certain groups, the resulting model inherits and perpetuates these biases. For instance, studies showed that Amazon's hiring algorithm was trained on past recruitment data, which primarily included male candidates (as women were historically not selected by human recruiters). As a result, the algorithm disfavoured female applicants and thereby reinforced existing gender biases.[26] Similarly, in facial recognition technology, studies have shown that AI systems perform much better on white men than on Black women, leading to disparities in accuracy and potentially discriminatory outcomes.[27] Another example is the infamous COMPAS case, where Black defendants were more likely to get false positives of being classified as high-risk for recidivism than white defendants, due to historical data that was biased against Black people.[28] A last example concerns large language models (LLMs) of which a UNESCO study has found that they exhibit gender biases, often favouring male over female representations in text generation tasks.[29]

Algorithmic discrimination is difficult to identify and address due to the complexity and opacity of AI models. Many AI models function as black boxes, making it difficult to interpret their decision-making processes and assess whether they have been trained fairly and representatively. While efforts may be made to use unbiased data, truly bias-free datasets are extremely rare in practice. Bias can emerge at different stages, such as during data collection, preprocessing and model development. For instance, labelling and annotating data often involve human subjectivity, which may introduce unintentional biases. Poor selection and preparation of data can also lead to discrimination.

Several methods have been proposed to help **detect and mitigate biases**.[30] One approach involves conducting algorithmic audits, which require full access to the underlying code and data to examine whether the AI systems processed data in a way that leads to discrimination. Another method involves creating synthetic profiles to evaluate whether AI-driven decision would vary based on protected

[25] FRA, 2020, *Getting the future right – Artificial intelligence and fundamental rights,* https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf.
[26] Dastin, J., 2018, *Insight - Amazon scraps secret AI recruiting tool that showed bias against women*, https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/.
[27] Buolamwini, J. & Gebru, T., 2018, 'Gender shades: Intersectional accuracy disparities in commercial gender classification', *Conference on Fairness, Accountability, and Transparency*, pp. 1-15.
[28] Larson, J., Mattu, S., Kirchner, L. & Angwin, J., 2016, 'How We Analyzed the COMPAS Recidivism Algorithm', *ProPublica*.
[29] UNESCO, 2024, *Challenging systematic prejudices: an investigation into bias against women and girls in large language models*, https://unesdoc.unesco.org/ark:/48223/pf0000388971.
[30] FRA, 2018, *#BigData: Discrimination in data-supported decision making*, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-focus-big-data_en.pdf.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

characteristics. For example, in a hiring algorithm, one could submit two identical applications that differ only in gender and assess whether outcomes change. Greater transparency in AI systems would also facilitate bias detection, allowing scrutiny of data, algorithms and decision-making processes. However, full transparency is often not feasible due to concerns related to intellectual property rights, national security and proprietary data protection.[31]

Within ACHILLES, different systematic approaches are considered. Besides standard group fairness approaches, Task 1.2 advances research on counterfactual fairness, which evaluates whether an AI system's decisions would remain the same if a protected characteristic is altered. Data-centric methodologies involve auditing datasets to assess their distribution, representativity, bias and correctness before model training. A strong emphasis is put on ensuring high-quality data, and techniques like adversarial training are considered to help minimise disparities in model outputs by adjusting how different groups are treated in decision-making processes. Moreover,

Task 4.4 will conduct research to ascertain how the technology developed in ACHILLES aligns with the values of the various stakeholders and how the designed artefacts affect human decision-making. The results will be compiled in D4.6 "*Socio-technical scenarios and interaction design*" (M12), led by FhAICOS.

3.6 (Intellectual) property

Intellectual property is recognised as a fundamental right under Art. 1 ECHR, Protocol No. 1 to the ECHR[32] and Art. 17(2) CFR. Art. 1 ECHR establishes that every natural and legal person is entitled to the peaceful enjoyment of his possessions, and shall not be deprived of possessions, unless in the public interest and subject to conditions by law. Art. 17(2) CFR explicitly states that intellectual property shall be protected. This protection extends not only to literary and artistic property but also patent, trademark rights and associated rights.

The intersection of AI and intellectual property raises different legal and ethical challenges, such as who holds ownership rights over AI-generated output, whether such output qualifies for intellectual property protection, and the extent to which AI systems can be trained on copyright-protected materials.[33] As such, this right is especially relevant for ACHILLES IDE's outcomes as well as the SCRIPTA and HERA use cases.

3.7 Healthcare

[31] FRA, 2019, *Data quality and artificial intelligence – mitigating bias and error to protect fundamental rights*, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2019-data-quality-and-ai_en.pdf.
[32] ECtHR No. 73049/01, Anheuser-Busch Inc. v. Portugal, 11 January 2007; ECtHR, 2024, *Guide on Article 1 of Protocol No. 1 to the European Convention on Human Rights – Protection of property*. https://ks.echr.coe.int/documents/d/echr-ks/guide_art_1_protocol_1_eng.
[33] Vanherpe, J., 2024, 'Artificial Intelligence and Intellectual Property Law', *The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence*, pp. 211-227.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

The right to healthcare has been recognised in Art. 35 CFR and Art. 12 ICESCR. It entails that everyone has the right to access healthcare and the right to benefit from medical treatment. Healthcare is grounded in respect for human dignity and the intrinsic worth of every individual, and is usually linked to principles like autonomy, transparency, accountability or inclusivity.[34]

However, AI-driven healthcare solutions may present both ethical and legal challenges to this right and principles.[35] For instance, too much delegation or reliance on medical AI systems could compromise **human autonomy** and urge the need for individuals to retain control over medical choices and healthcare systems. Medical AI solutions also necessitate more transparency and explainability to improve the overall system quality and protect patient safety and public health. It also raises difficult questions regarding accountability and responsibility when many actors are involved. Linked to sub-section 3.5, it is important that the use of medical AI systems does not result in discriminatory outcomes. Therefore, developers should ensure that AI (training) data is free from biases, accurate, diverse and representative.

The right to healthcare is specifically relevant for ACHILLES as two of the use cases, HERA and the one related to ophthalmologic diagnostic support, take place in the context of healthcare. Therefore, it will be important to take both ethical and legal considerations into account for piloting the HERA and healthcare use cases during ACHILLES.

3.8 Environmental protection

Art. 37 CFR and Art. 12(2)(b) ICESCR recognise environmental protection as a fundamental right. While environmental concerns have long been researched, the rise of AI has brought renewed attention to the need for sustainable practices in technology development.[36] The growing emphasis on environmental protection and sustainability in the context of AI stems from the **ecological impact of AI systems**, such as electronic waste, water consumption, reliance on critical minerals and rare elements and high electricity usage – especially in the case of more sophisticated models (i.e. generative AI) that require even more resources.[37]

Within ACHILLES, sustainability is a key consideration. It is important to integrate sustainable practices both during the development of AI technologies within the project as well as the AI systems and the IDE platform.

---

[34] Mennella, C., Maniscalco, U., De Pietro, G. & Esposito, M., 2024, 'Ethical and regulatory challenges of AI technologies in healthcare: A narrative review', *Heliyon*.
[35] Mennella, C., Maniscalco, U., De Pietro, G. & Esposito, M., 2024, 'Ethical and regulatory challenges of AI technologies in healthcare: A narrative review', *Heliyon*.
[36] UN, 2024, *AI has an environmental problem. Here's what the world can do about that*, https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about.
[37] UN, 2024, *Artificial intelligence (AI) end-to-end: The environmental impact of the full AI life cycle needs to be comprehensively assessed*, https://wedocs.unep.org/bitstream/handle/20.500.11822/46288/AI-Environmental-Impact-Issues-Note.pdf?sequence=3&isAllowed=y.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

4 ARTIFICIAL INTELLIGENCE

4.1 European AI Act

4.1.1 Introduction

Since the main objective of ACHILLES is to develop an IDE that provides AI developers with relevant compliance information and technical support, it is essential to assess the applicable AI requirements for the project and its various activities. In recent years, several legal initiatives have emerged. One of them is the European AI Act.[38] The AI Act promotes human-centric, trustworthy and sustainable AI, and respect for fundamental rights.

While the AI Act will officially apply from 2 August 2026, two years after its entry into force, certain provisions take effect earlier. From 2 February 2025, provisions regarding the prohibited AI practices became applicable. From 2 August 2025, provisions regarding general purpose AI (GPAI) models are applicable. The provisions regarding high-risk AI systems will not take effect until 2 August 2027, three years after entry into force (Art. 113). Not all provisions of the AI Act are already applicable, but will enter into force during the project's timeline, the ACHILLES consortium adopts a forward-looking approach to demonstrate its commitment to align the project's activities with the obligations in the AI Act. This ensures the AI systems designed, developed and used in the project – even afterwards – are and will remain sustainable.

After providing an overview of the AI Act's scope and key definitions, the analysis will focus on a selection of obligations most relevant to the project.

4.1.2 Scope and key definitions

The AI Act establishes obligations for the placing on the market, putting into service and the use of AI systems in the EU. It is the first horizontal framework in the EU setting out rules towards AI developers, providers and users to ensure AI systems are safe and respect fundamental rights and EU values.[39]

In terms of **territorial scope**, the AI Act applies to (a) providers placing on the market or putting into service AI systems or general-purpose AI models in the Union, (b) deployers of AI systems that have their place of establishment or are located within the Union, (c) providers and deployers of AI systems that have their place of establishment or are located in a third country, where the output produced by the AI system is used in the Union, (d) importers and distributors of AI systems, (e) product manufacturers placing on their market or putting into service AI systems together with their product and

[38] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 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 (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024.

[39] EU, "European approach to artificial intelligence", 9 December 2024, https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

under their own name or trademark, (f) authorised representatives of providers, which are not established in the Union, and (g) affected persons that are located in the Union (Art. 2). However, certain areas are excluded from the AI Act's scope, including areas outside the scope of EU law (such as military, defence, and national security). Exemptions also apply to AI systems with the sole purpose of scientific research and development (R&D), pre-market research, testing or development activity of AI systems/models, AI deployers who are natural persons using AI systems purely for non-professional activity, or AI systems released under free and open-source licences – unless such systems are placed on the market or put into service as prohibited, high-risk or limited risk AI systems (Art. 2).

Regarding the **personal scope**, the AI Act is applicable to AI systems, defined 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*" (Art. 3(1)(1)).

As for ACHILLES, all project partners are established in the EU, except CUOMOIT. However, Art. 2(1)(c) clarifies that when the output produced by an AI system is used in the EU, the AI Act is applicable even though the provider or deployer is not established or located in the EU. Given this, it is likely that the AI Act applies – in territorial terms – to the ACHILLES project activities. In addition, as for the personal scope, the ACHILLES IDE is likely to qualify as an AI system. The IDE is a machine-based system designed to operate with varying levels of autonomy, referring that they operate with some degree of independence from human involvement (recital 12), and can infer from input it receives how to generate outputs, such as recommendations that can influence both physical and virtual environments. Indeed, as a virtual assistant (a copilot) integrated into the user interface, the IDE will streamline AI development processes by providing real-time support, including project navigation assistance, best practice suggestions, and easy access to relevant technical, ethical, and legal documentation. It will ensure that developers have a conversational partner throughout the AI development lifecycle, from co-design to monitoring after deployment. The IDE will integrate plugins adaptable to various domains, and will learn from data how to achieve this objective. Moreover, the four use cases envisaged within ACHILLES (SCRIPTA, HERA, healthcare and identify verification) are likely to qualify as an AI system as well. In the future, AI systems developed through the IDE should always be assessed to determine whether they fall under the scope of the AI Act.[40]

The AI Act adopts a **risk-based approach**. This means that the obligations vary based on the risk posed by an AI system. Traditionally, four risk levels are distinguished:

[40] In February 2025, the European Commission (EC) published it Guidelines on AI system definition to facilitate the AI Act's rules application.
European Commission, 2025, *The Commission publishes guidelines on AI system definition to facilitate the first AI Act's rules application*, https://digital-strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

- Unacceptable risk (prohibited)
- High risk (permitted but subject to obligations)
- Limited risk (transparency obligations)
- Minimal risk (allowed without further due)

The AI Act prohibits AI systems deemed to pose an **unacceptable risk** to fundamental rights (Art. 5). Prohibited practices are: (a) AI systems using subliminal techniques, (b) AI systems exploiting vulnerabilities of persons, (c) social scoring systems, (d) AI systems for risk assessments, (e) AI systems for compiling facial recognition databases, (f) emotion recognition systems in workplace or education, (g) biometric categorisation systems, (h) real-time remote biometric identification systems in publicly accessible spaces for law enforcement.

AI systems with a **high risk** are not prohibited but subject to mandatory requirements and ex-ante conformity assessments, as they are deemed to pose a high risk to the health, safety or fundamental rights of individuals (Art. 6 AI Act and Chapter III, Section 2). To qualify as high-risk, an AI system must meet two cumulative criteria: (1) the AI system is intended to be used as a safety component of a product or the AI system is itself a product covered by Union legislation listed in Annex I,[41] and (2) that product is required to undergo a third-party conformity assessment. In addition, Annex III classifies AI systems as high-risk when used in specific areas: (1) biometrics; (2) critical infrastructure; (3) education and vocational training; (4) employment, workers management and access to self-employment; (5) access to and enjoyment of essential private services and essential public services and benefits; (6) law enforcement; (7) migration, asylum and border control management; and (7) administration of justice and democratic processes. Nonetheless, Art. 6(3) allows for exceptions: an AI system listed in Annex III may be exempted from being a high-risk AI system if it does not pose a significant risk of harm to the health, safety or fundamental rights of natural persons[42] – which is to be considered by the provider of the AI system itself. Notably, AI systems used to profile natural persons are always considered high-risk.

Limited risk AI systems are systems that pose a low risk. Art. 50 states that AI systems intended to directly interact with natural persons must comply with minimum transparency obligations. The AI systems must be designed and developed so that natural persons are aware that they are interacting with an AI system, unless this is clear the viewpoint of the user, taking into account the circumstances and context of use.

[41] Annex I include, for instance, union harmonisation legislation on machinery, safety of toys, lifts, radio equipment, medical devices, in vitro diagnostic medical devices, but also civil aviation security, marine equipment, rail systems and so on.
[42] Art. 6(3) AI Act specifies that AI systems referred to in Annex III shall not be considered to be high-risk when AI systems is intended (a) to perform a narrow procedural task, (b) to improve the result of a previously completed human activity, (c) to detect decision-making patterns or deviations from prior decision-making patterns and not meant to replace or influence previously completed human assessment without proper human review, or (d) to perform a preparatory task to an assessment relevant for purposes of use cases listed in Annex III.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Minimal or no risk AI systems, such as spam filters in a mailbox, are permitted without additional obligations. AI providers may voluntarily conform to the requirements imposed on high-risk AI systems or adopt voluntary codes of conduct.

The AI Act also establishes specific obligations on providers of **GPAI models** and **GPAI systems**.[43] A GPAI model is defined as an "*AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market*" (Art. 3(63)). A GPAI system is defined as "*an AI system which is based on a general-purpose AI model and which has the capability to serve a variety of purposes, both for direct use as well as for integration in other AI systems*" (Art. 3(66)). The AI Act distinguishes between 'traditional' GPAI models and GPAI models posing a **systemic risk**. The latter refers to "*risks that are specific to the high-impact capabilities of general-purpose AI models, having a significant impact on the Union market due to their reach, or due to actual or reasonably foreseeable negative effects on public health, safety, public security, fundamental rights, or the society as a whole, that can be propagated at scale across the value chain*" (Art. 3(65)). Whereas providers of 'traditional' GPAI models are subject to documentation keeping, copyright protection and transparency obligations (Art. 53), providers of GPAI models with a systemic risk face stricter additional requirements, such as model evaluations, risk mitigation measures, documentation keeping, and cybersecurity protection.

The AI Act's regulatory framework has traditionally been visualised as a pyramid, as shown in Figure 2. However, with the existence of GPAI models and systems, this traditional structure may no longer fully capture the entire regulatory framework. Therefore, Prof dr Palmiotto proposes an alternative representation, as shown in Figure 3.

[43] Please note, these obligations will only apply one year after the AI Act's entry into force, i.e. mid-2025.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

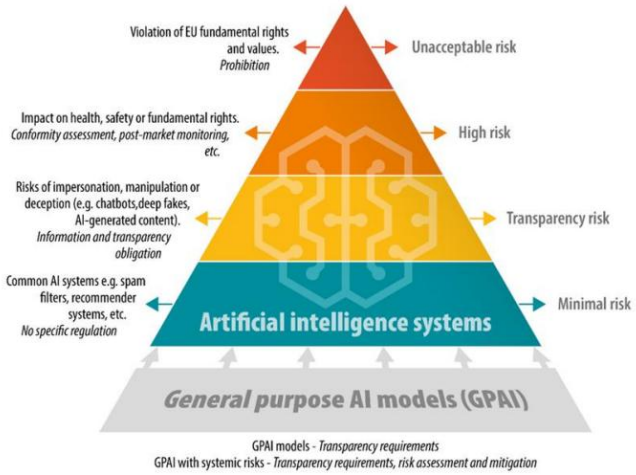

Figure 2. The EU AI Act risk-based approach (European Union 2024)44

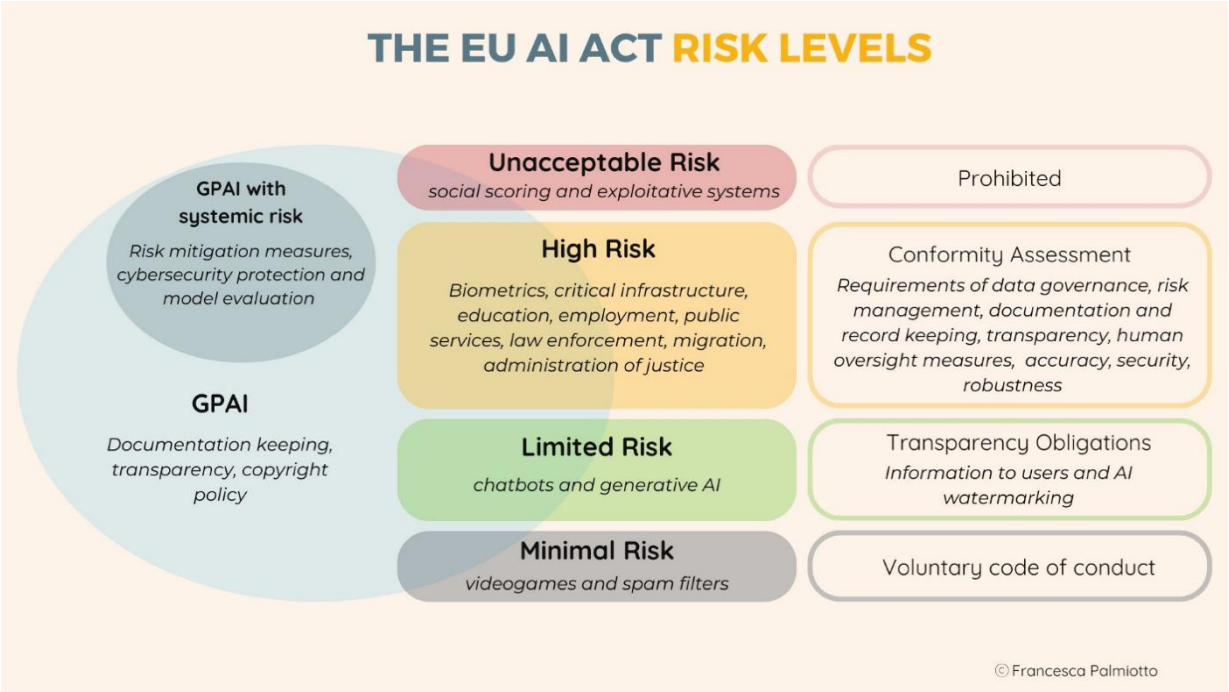

Figure 3. The EU AI Act Risk Levels, by Francesca Palmiotto (December 2024).

The next step of ACHILLES involves assessing how the IDE qualifies under the AI Act – if so. First, the IDE does not seem to qualify as a GPAI model as defined in Art. 3(63). Whereas it is likely to presume that the IDE is trained on a large amount of data, it seems unlikely to display a significant generality and

44 European Union, 2024, *Artificial intelligence act*, https://www.europarl.europa.eu/RegData/etudes/BRIE/2021/698792/EPRS_BRI(2021)698792_EN.pdf.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

capability of performing a wide range of tasks. The IDE is designed for a single specific purpose, namely, to function as a copilot that provides real-time support, including project navigation assistance, best practice suggestions, and easy access to relevant technical, ethical, and legal documentation. The IDE also appears to not classify as a GPAI system under Art. 3(66). This would require that the IDE copilot is both based on a GPAI model, such as GPT-4, and serves a variety of purposes. As mentioned, the IDE copilot appears to be designed for one specific purpose, and the second condition is thus not fulfilled. Nevertheless, the final qualification will ultimately depend on technical considerations that will become clearer as the project progresses.

In addition, the IDE does not seem to qualify as a high-risk AI system under Art. 6 as it neither falls under Union harmonisation legislation nor any of the areas mentioned in Annex III. However, the IDE is likely to qualify as an AI system posing limited risk and therefore requiring compliance with **transparency** obligations. Art. 50(1) stipulates that AI systems intended to interact with natural persons must be designed and developed in a way that the users are informed that they are interacting with an AI system, unless this is obvious – taking into account the circumstances and context of use.[45] This aligns with the IDE's core function as a conversational partner for AI system developers throughout the AI development lifecycle.

In addition to the qualification of the IDE, it will be important that for each AI system being developed through the IDE, it will be able to identify the type of AI system it concerns, as classifications trigger different obligations. As for the use cases within the project, it seems that SCRIPTA and HERA would trigger transparency obligations under Art. 50(1). Given the high thresholds for qualifying as a GPAI model and/or system, it seems unlikely they would be qualified as such. For the HERA and healthcare use cases, it must be assessed whether they fall under high-risk AI systems subject to Union harmonisation legislation listed in Annex I (art. 6(1)).[46] Annex I, Part A includes the Medical Devices Regulation (*infra* 9.1), meaning that AI systems used in healthcare may fall within this category. As the use case validation WP progresses and more details about the use cases become available, a detailed assessment must be conducted to determine whether these AI systems trigger the application of Union harmonisation legislation. If so, they will be considered high-risk AI systems, requiring compliance with the corresponding obligations. Similarly, for the identity verification use case, it should be determined whether it can be qualified as a high-risk AI system by triggering Art. 6(2) j Annex III point 1(a), which covers AI systems used in the area of biometrics. At this stage, it appears unlikely that the use case would fall within this classification. Its primary function is limited to comparing images rather than engaging in remote biometric identification. Nevertheless, this assessment should be evaluated once more detailed information about the use case becomes available. Moreover, even if the use case would fall under Annex III, point 1(a), the exclusion in the second paragraph would likely apply. This exclusion states that "*AI systems intended to be used for biometric verification the sole purpose of which is to*

[45] See also recital 132.
[46] van Kolfschooten, H. & van Oirschot, J., 'The EU Artificial Intelligence Act (2024): Implications for healthcare', *Health Policy*.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

confirm that a specific natural person is the person he or she claims to be" are not classified as high-risk. Since this aligns with the purpose of the identity verification use case, it strongly suggests that the use case would be excluded from the high-risk category under the AI Act.

Most of the provisions in the AI Act apply to **providers**, defined as natural or legal persons, public authorities, agencies or other bodies that develop an AI system or a GPAI model or that have the AI systems or GPAI model developed and place it on the market or put it into service under their own name or trademark, whether for payment or free of charge (Art. 3(3)). Conversely, **deployers** are defined as a natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity (Art. 3(4)). As for ACHILLES, it will be important to identify the provider(s) and/or deployer(s) for each AI system developed, as they will be responsible for compliance with the AI Act.

The following sections provide an overview of the key obligations to consider when developing, deploying and using AI systems. It focuses on the requirements for high-risk AI systems while also offering a broader overview of other obligations. In addition, the analysis outlines requirements relevant to the energy consumption of AI systems, the importance of fundamental rights impact assessments, and transparency obligations for certain AI systems.

4.1.3 Requirements for high-risk AI systems

4.1.3.1 Risk management system

According to Art. 9 AI Act, a risk management system shall be established, implemented, documented and maintained in relation to the high-risk AI system. A risk management system is a continuous iterative process that goes throughout the entire lifecycle of high-risk AI systems and requires regular review and updates. First, it identifies and analyses the reasonably foreseeable risks of the AI system towards health, safety or fundamental rights, evaluation of these risks, evaluation of other risks possibly arising, and which mitigation measures are adopted to target risks. To identify the most appropriate and targeted mitigation measures, Art. 9(6) requires the system to be tested to see whether it performs consistently with the intended purpose. Testing can occur throughout the systems' development process, but in any case, always before being placed on the market or put into service.

4.1.3.2 Data and data governance

The data and data governance requirements set out in Art. 10 are of much importance to ACHILLES. Many project activities and AI systems involve (personal) data to be trained upon. Besides the requirements for personal data (*infra* 5) and data governance (*infra* 6), the AI Act also includes data and data governance requirements in the sense that high-risk AI systems training their AI models with data (whether personal or not) shall be developed on the basis of training, validation and testing data sets that meet quality criteria and are subject to data governance and management practices appropriate for the intended purpose of the high-risk AI system. Those practices are:

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

(a) the relevant *design choices*;
(b) *data collection processes* and the *origin of data*, and in the case of personal data, the original purpose of the data collection;
(c) relevant *data-preparation processing operations*, such as annotation, labelling, cleaning, updating, enrichment and aggregation;
(d) the formulation of *assumptions*, in particular with respect to the information that the data are supposed to measure and represent;
(e) an assessment of the *availability, quantity and suitability* of the data sets that are needed;
(f) examination in view of possible *biases* that are likely to affect the health and safety of persons, have a negative impact on fundamental rights or lead to discrimination prohibited under Union law, especially where data outputs influence inputs for future operations;
(g) appropriate *measures* to detect, prevent and mitigate possible biases identified according to point (f);
(h) the identification of relevant *data gaps or shortcomings* that prevent compliance with this Regulation, and how those gaps and shortcomings can be addressed.

It is furthermore required that training, validation and testing data are relevant, sufficiently representative and, to the best extent possible free of errors and complete.

Based on these requirements, providers of high-risk AI systems must **identify any potential bias** in their data and data governance practices and take appropriate measures to mitigate them to avoid a negative impact on fundamental rights, especially the right to non-discrimination. For this purpose, paragraph 5 goes a step further and allows, for the purpose of ensuring bias detection and correction, that special categories of personal data as understood under the GDPR can be processed when safeguards are taken to protect fundamental rights and freedoms of natural persons. This is because research has shown that such use of sensitive data can help detect discrimination.[47] Although discussed in more detail down below (*infra* 5.1.5), sensitive data in the sense of the GDPR refers to racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union memberships, genetic data, biometric data, health data concerning a person's sex life or sexual orientation (Art. 9(1) GDPR). Studies indicate that without sensitive data, providers may struggle to test proxies in data. For example, suppose an AI system is used to make decisions that could result in indirect discrimination due to a certain ethnicity. To identify potential biases in its decision-making process, the developer must analyse whether ethnicity influences the outcomes. However, simply removing explicit references to ethnicity from the AI system's design will not eliminate bias. As noted earlier (*supra* 3.5), certain attributes (while not direct protective grounds or sensitive data) can serve as proxies for ethnicity, such as postal code. To

[47] van Bekkum, M. & Borgesius F.Z., 2023, 'Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception?', *Computer Law & Security Review*; van Bekkum, M., 'Using sensitive data to de-bias AI systems: Article 10(5) of the EU AI act', *Computer Law & Security Review*.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

determine whether an attribute functions as a proxy, providers must collect and analyse ethnicity data.[48]

Therefore, Art. 10(5) AI Act allows providers of high-risk AI systems to process sensitive data specifically **for bias detection and mitigation**. Nevertheless, these provisions should not be seen as a blanket exemption from GDPR compliance. While the AI Act allows such processing, the GDPR still applies, and both frameworks – although potentially in tension – must be reconciled. Under Art. 9 GDPR, the processing of special categories of data is, in principle, prohibited unless a specific legal basis applies. One such ground is the explicit consent of the data subject. Therefore, when applying Art. 10(5) AI Act, providers must ensure that a valid legal basis under the GDPR is met. Some have argued that Art. 9(2)(g) GDPR could be invoked in this context, which allows the processing of sensitive data when necessary for substantial public interest reasons since fighting discrimination could be seen as a relevant public interest.[49] However, even in that case, the rights of the data subjects must remain protected, and providers must take appropriate measures to safeguard other fundamental rights and interests. The principles of the GDPR, such as data minimisation, purpose and storage limitation, are still applicable and robust cybersecurity measures should be taken to prevent data leaks.[50] Hence, although Art. 10(5) AI Act provides a legal basis for processing sensitive data for detecting and mitigating biases, it must be interpreted narrowly and only applied when strictly necessary to protect fundamental rights, such as non-discrimination.

For the sake of completeness, as mentioned, discrimination can also arise from processing non-sensitive personal data, such as postal code, and people's height or weight. Even when such personal data is being used for the detection and mitigation of biases, the GDPR remains fully applicable. This means that a valid legal ground under Art. 6 GDPR should be respected (*infra* 5.1.4).

4.1.3.3 Technical documentation and record-keeping

Art. 11 and 12 AI Act require to drawn up technical documentation of the high-risk AI system and keep automatic recording of events (logs), respectively. These transparency and accountability-enhancing requirements ensure comprehensive information is kept on how high-risk AI systems have been developed and how they perform. System traceability enables to verify compliance with requirements under the AI Act and facilitates post-market monitoring.

[48] van Bekkum, M., 'Using sensitive data to de-bias AI systems: Article 10(5) of the EU AI act', *Computer Law & Security Review*.
[49] van Bekkum, M., 'Using sensitive data to de-bias AI systems: Article 10(5) of the EU AI act', *Computer Law & Security Review*.
[50] CJEU C-524/06 Heinz Huber v. Germany, 16 December 2008, ECLI:EU:C:2008:724; European Parliament, 2025, *Algorithmic discrimination under the AI Act and the GDPR*, https://www.europarl.europa.eu/RegData/etudes/ATAG/2025/769509/EPRS_ATA(2025)769509_EN.pdf.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

The technical documentation is prepared to demonstrate that the high-risk AI system complies with the requirements of the AI Act and should include at least the elements set out in Annex IV.[51] In case a high-risk AI system constitutes a product covered by Union harmonisation legislation listed in Section A of Annex I, a single set of technical documentation can be drawn up that contains both the information required in the Annex IV AI Act and the information required in the harmonisation legislation. For instance, when the Medical Device Regulation (*infra* 9.1) – falling under Section 1 of Annex I – already demands for technical documentation to be drafted, the information required in Annex IV of the AI Act can be integrated into the same document.

As for the record-keeping of events (logs), these should be recorded when relevant to identify the system presenting a risk or a substantial modification, to facilitate post-market monitoring, and to monitor the operation of high-risk AI systems. Specifically with regard to remote biometric identification systems referred to in point 1(a) Annex III, logging must also provide (a) a record of the period of each use of the system (start date and time and end date and time of each use), (b) the reference database against which input data has been checked by the system, (c) the input data for which the search has led to a match, and (d) the identification of the natural persons involved in the verification of the results, as referred to in Art. 14(5). Art. 19 AI Act further clarifies that it should be the providers that keep the logs as referred to in Art. 12(1) AI Act when such logs are under their control.

In addition, Art. 18 AI Act requires providers of high-risk AI systems to keep at their disposal certain documentation, including technical documentation referred to in Art. 11, documentation regarding the quality management system referred to in Art. 17, documentation regarding changes approved by notified bodies if applicable, decisions and other documents issued by notified bodies, and EU declaration of conformity as referred to in Art. 47.

4.1.3.4 Transparency and provision of information to deployers

Art. 13 AI Act imposes a set of transparency and information obligations. In the first place, it is required that high-risk AI systems – given the concerns about the opacity and complexity and the high-risk area they are used in – should be designed and developed in a way that ensures that their operation is sufficiently transparent to enable deployers to interpret the system's output and use it appropriately. This system-level transparency does not necessarily imply full model interpretability or explainability

[51] Technical documentation in Art. 11(1) AI Act should contain at least the following information: (1) a general description of the AI system, a detailed description of the AI system and the process for its development, (3) detailed information about the monitoring, functioning and control of the AI system, (4) a description of the appropriateness of the performance metrics for the specific AI system, (5) a detailed description of the risk management system in accordance with Art. 9, (6) a description of relevant changes made by the provider to the system through its lifecycle, (7) a list of the harmonised standards applied and where no such harmonised standards have been applied, a detailed description of the solutions adopted to meet the requirements set out in Chapter III, Section 2, including a list of other relevant standards and technical specifications applied, (8) a copy of the EU declaration of conformity referred to in Art. 47, (9) a detailed description of the system in place to evaluate the AI system performance in the post-market phase in accordance with Art. 72.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

(e.g. being able to trace every decision patch in a neural network), but rather focuses on ensuring that the deployers understand the system's intended functions, capabilities and limitations (recital 72). Deployers should be able to assess how the system works in practice, its strengths and limitations, and how to use it responsibly. Therefore, paragraphs 2 and 3 specify that providers must issue instructions for use that include concise, complete, correct and clear information that is relevant, accessible and comprehensible to deployers. Providers could, for instance, include illustrative examples to make information more comprehensible for deployers, and adapt instructions based on target deployers (recital 72). The instructions should at least contain information about the identity and contact details of the provider; the characteristics, capabilities and limitations of performance of the system; changes to the system and its performance, if any; human oversight measures taken; computational and hardware resources needed; expected system lifetime and maintenance needs; and description of mechanisms to store and interpret logs. These instructions should help deployers to use the system, support informed decision-making, and ensure they are correctly educated about the use of the AI system.

4.1.3.5 Human oversight

Art. 14 AI Act imposes that high-risk AI systems must be designed and developed in a way that they can be effectively overseen by natural persons when they are in use. This is to ensure that deployers can **oversee the system's functioning and impacts over the lifecycle** (recital 73) – and thus goes beyond the design and development phase. Human oversight ensures that systems are used for their intended purpose(s), prevent risks, and that humans overseeing have the required knowledge to understand the capabilities and limitations of systems. To ensure human oversight, oversight measures are to be taken, either identified or built into the high-risk AI systems by the provider itself before placing them on the market or putting them into service, or measures identified by providers but implemented by deployers. Recital 73 specifies that such measures could be, for instance, that a system is subject to in-built operational constraints that cannot be overridden by the system itself and is responsive to a human operator, and that persons assigned human oversight have the necessary competence, training and authority to carry out that role.

Specifically with regard to remote biometric identification systems referred to in point 1(a) Annex III, the measures must ensure that, in addition, no action or decision is taken by the deployer based on the system's identification, unless that identification has been separately verified and confirmed by at least two natural persons with the necessary competence, training and authority – these people can be from one or more entities and include persons operating or using the system. Recital 73 clarifies that such additional measures are required due to potential consequences in case of an incorrect match. It further clarifies that this additional requirement should not pose an unnecessary burden or delay, and sometimes that separate verifications by the different persons are automatically recorded in the logs generated by the system.

Therefore, paragraph 4 specifies that the high-risk AI systems are to be provided to the deployer so that the persons who are doing human oversight are enabled (a) to properly understand the capacities and

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

limitations of the system and monitor its operation, in order to detect and address anomalies, dysfunctions and unexpected performance, (b) to remain aware of automation bias, in particular when the system provides information or recommendations, (c) to interpret the system's output, (d) to decide not to use the system or to otherwise disregard, override or reverse its output, (e) to intervene in the system's operation or interrupt the system through a 'stop' button or a similar procedure that allows the system to safely come to a halt. Hence, the provider has some flexibility in choosing to implement it appropriately and proportionately.[52]

While human oversight is a key obligation, its practical implementation poses some challenges. Two main concerns have been identified. First, the concern regarding oversight bottleneck arises, referring to the fact that humans struggle to effectively monitor or intervene in complex AI decision-making. Many complex AI systems are often characterised by their black-box decision-making, making it impossible for humans to really understand their reasoning. However, without a clear understanding, human intervention becomes challenging. A second concern relates to scalable oversight, which refers to effectively monitoring and controlling AI systems, especially as they become more complex and widely adopted. This issue is further exacerbated by the risk of automation bias, where humans tend to trust and follow AI recommendations simply because they assume AI is more objective and trustworthy. As a result, human overseers are prone to automation bias, making human oversight an imperfect safeguard. Given these limitations, human oversight alone is not sufficient. Instead, all other requirements must be read **in tandem**.[53]

4.1.3.6 Accuracy, robustness and cybersecurity

Art. 15 AI Act states that high-risk AI systems must be designed and developed in a way that they have appropriate levels of accuracy, robustness and cybersecurity, and that they perform consistently so throughout their lifecycle. It implies that systems must be **resilient** against errors, faults or inconsistencies, and for that reason, **technical and organisational measures** must be taken – taking into account relevant circumstances and risks. Recital 74 clarifies that in the instructions of use in which the expected level of performance metrics is declared, providers are urged to communicate that information to deployers in a clear and easily understandable way.

Paragraph 5 further emphasises the need to ensure robustness by implementing **technical redundancy solutions**, such as backup or fail-safe plans. Technical robustness is indeed important to avoid harmful and undesirable behaviour resulting from the system. Fail-safe plans could include mechanisms to safely interrupt the system's operation (recital 75). Systems must also be **resilient**

[52] Fink, M., *Human Oversight under Article 14 of the EU AI Act*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5147196.
[53] Fink, M., *Human Oversight under Article 14 of the EU AI Act*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5147196, H.K., 2025, *The Limits of Human Oversight: What Alignment Research Reveals About the EU AI Act's Gaps*, https://www.linkedin.com/pulse/limits-human-oversight-what-alignment-research-eu-ai-acts-hernandez-zlclf/; Jarovsky, L. *Can human really oversee AI?,* https://www.luizasnewsletter.com/p/can-humans-really-oversee-ai.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

against attempts by unauthorised third parties to alter the system's use, outputs or performance by exploiting vulnerabilities, and thus cybersecurity is highly important for systems. Technical solutions shall include appropriate measures to prevent, detect, respond to, resolve and control for attacks trying to manipulate the training data (i.e. data poisoning), or pre-trained components used in training (i.e. model poisoning), inputs designed to cause the AI model to make a mistake (i.e. adversarial examples or model evasion), confidentiality attacks or model flaws. To ensure an appropriate level of cybersecurity, providers of high-risk AI systems must implement suitable measures, such as security controls, while taking into account the underlying ICT infrastructure where relevant (recital 76).

Please note that sector-specific legislation related to cybersecurity complements Art. 15 AI Act and must be respected as well (*infra* 8).

4.1.3.7 Quality management system

According to Art. 17 and Art. 16(c) AI Act, providers of high-risk AI systems must put in place a quality management system that ensures compliance with the AI Act, drawing up the relevant documentation and establishing a robust post-market monitoring system (see also recital 81). The quality monitoring system must include at least twelve aspects as described in Art. 17: (a) a strategy for regulatory compliance, (b) techniques, procedures and systematic actions for the design, design control and design verification of the system, (c) techniques for the development, quality control and quality assurance of the system, (d) examination, test and validation procedures, (e) technical specifications next to harmonised standards, (f) systems for data management, including data acquisition, collection, analysis, labelling, storage, filtration, mining, aggregation, retention and so on, (g) risk management system as referred to in Art. 9 AI Act, (h) setting-up, implementation and maintenance of post-marketing system, (i) procedures relating to reporting serious incidents, (j) handling of communication with national competent authorities, (k) procedures for record-keeping of all documentation and information, (l) resource management, and (m) accountability framework setting out responsibilities of management and staff regarding all aspects in list.

4.1.3.8 Obligations of deployers of high-risk AI systems

Art. 26 AI Act sets forward obligations towards **deployers** of high-risk AI systems. It starts by outlining that deployers must take **appropriate technical and organisational measures** to ensure that the high-risk AI systems they deploy are **in accordance with instructions for use**.

Deployers shall assign human overseers to comply with their obligations. Deployers who exercise control over input data must ensure that input data is relevant and representative in view of the intended purpose of the system.

Deployers must monitor the operations of the AI system based on instructions for use. When they believe the operation might result in risks that can adversely affect the health, safety or fundamental rights of persons, they must **inform** the provider or distributor and relevant market surveillance authority of the risk and suspend the use of the system. Deployers must keep **logs** automatically

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

generated by high-risk AI systems to the extent such logs are under their control, and for at least six months.

Importantly, Art. 25 clarifies that deployers, distributors, importers or any other third-party will be considered a provider of a high-risk AI system and thus subject to the requirements under Art. 16 when they either put their name or trademark on the system already on the market or put into service, either make substantial modifications to the system already placed on the market or put into service in such a way that it remains a high-risk system, or they modify the intended purpose of the system previously not classified as high-risk in such a way that it becomes a high-risk AI system.

Art. 25, paragraph 3, further clarifies that when high-risk AI systems are safety components or products covered by **Union harmonisation legislation** listed in Section A of Annex I, the product manufacturer is to be considered the provider of the high-risk AI system and subject to all obligations under Art. 16 when either the high-risk AI system is placed on the market together with the product under the name or trademark of the product manufacturer or when the high-risk AI system is put into service under the name or trademark of the product manufacturer after the product has been placed on the market. This is relevant for the healthcare use case should it fall under the Medical Devices Regulation.

4.1.4 Fundamental Rights Impact Assessment

Art. 27 AI Act introduces a novel impact assessment within the AI context, namely the Fundamental Rights Impact Assessment (FRIA). As the name implies, a FRIA entails an assessment of how the use of high-risk AI systems can impact fundamental rights. According to recital 96, the FRIA requirement was introduced to ensure that **deployers** identify specific risks to fundamental rights and take appropriate measures. This aims to increase accountability of both public and private actors using AI technologies, prevent abuses and unintended consequences, and foster responsible AI. The FRIA enhances transparency, accountability, and ethical practices, building trust in AI technologies by evaluating potential risks associated with high-risk AI systems.

Prior to deploying a high-risk AI system as defined in Article 6(2) – except for high-risk AI systems used in critical infrastructure (which is not of relevance in ACHILLES) – deployers that are **bodies governed by public law** or **private entities providing public services**, and **deployers of high-risk AI systems referred to in points 5(b)(c) Annex III** (i.e. high-risk AI systems intended to be used to evaluate the creditworthiness or establish credit scores of natural persons, or intended to be used for risk assessment and pricing in relation to life and health insurance) shall perform a FRIA.

A FRIA requires deployers to perform an assessment that consists of (a) a description of the deployer's processes in which the system will be used in line with its intended purpose, (b) description of period of time in which system is intended to be used, (c) categories of natural persons and groups likely to be affected by its use, (e) description of implementation of human oversight measures, and (f) measures to be taken in case of risks.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

As noted by different scholars, the effectiveness of FRIAs relies on the information exchanges between providers and deployers.[54] Once the impact assessment is completed, the deployer must notify the relevant market surveillance authority. In addition, the FRIA must be updated if there are any changes to the system or its application.

Deployers may rely on previously conducted FRIAs or other existing assessments carried out by providers. In the case where a data protection impact assessment (DPIA) (*infra* 5.1.8) has already been carried out and covers some aspects of the FRIA, Art. 27, paragraph 4 clarifies that the FRIA complements the DPIA. For instance, while a DPIA may identify non-discrimination risks specifically related to personal data, the FRIA can extend the analysis to a broader scope, such as identifying non-discrimination risks related to algorithms behind the AI system.[55] The AI Office is expected to develop a template for a questionnaire and an automated tool to facilitate compliance with the FRIA obligations, but these resources are not yet available.

Given that ACHILLES partners do not appear to qualify as deployers that are governed by public law, private entities providing public services or deployers of high-risk AI systems used to evaluate the creditworthiness or establish credit scores of natural persons, or intended to be used for risk assessment and pricing in relation to life and health insurance, it is unlikely that a FRIA will be required within ACHILLES. Nevertheless, whenever AI systems are developed through the IDE, it will be important to assess whether these future systems may trigger this obligation.

4.1.5 Transparency obligations

Another obligation for ACHILLES concerns the transparency requirements set out in Art. 50 AI Act. This can be relevant not only for the SCRIPTA and HERA use cases but also for the ACHILLES IDE, which is a user-friendly platform designed to support developers during the AI lifecycle.

Art. 50 consists of different obligations. Art. 50(1) states that AI systems intended to directly interact with natural persons must be designed in a way that these persons are **informed** they are interacting with an AI system rather than a person, unless really obvious for a reasonably well-informed person. This means that providers should include information towards natural person users of their interactional systems, for instance, through the use of a pop-up message. This counts for both the SCRIPTA and HERA use cases, and the ACHILLES IDE.

Art. 50(2) further requires that providers of AI systems generating audio, image, video or text content (such as both the use cases and IDE) must ensure that the outputs of their system are marked in a

[54] Costentini, A. e.a., 2025, *Assessing the Impact of Artificial Intelligence Systems on Fundamental Rights*, https://www.medialaws.eu/wp-content/uploads/2025/03/Assessing-the-Impact-of-Artificial-Intelligence-Systems-on-Fundamental-Rights.pdf.

[55] FRA, 2020, *Getting the future right – Artificial intelligence and fundamental rights,* https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

machine-readable format and detectable as artificially generated or manipulated. Technical solutions must be effective, interoperable, robust and reliable.

Art. 50(3) specifically requires for biometric categorisation systems to inform natural persons exposed thereto about their operation. In addition, Art. 50(4), second paragraph, states that deployers of AI systems generating text published with the purpose of informing the public on matters of public interest must disclose that the text has been artificially generated or manipulated. However, this provision does not seem relevant to the project, as it does not pertain to text that is meant for the public interest.[56]

4.1.6 Energy consumption

As the ACHILLES project strives to develop sustainable AI systems by promoting less data-intrusive and energy-consuming models without compromising the models' quality, it also aims to focus on energy efficiency in the deployment stage. ACHILLES offers techniques to minimise energy consumption and optimise models for existing infrastructure. Art. 40(2) specifically states the importance of improving AI systems' resource performance, particularly by minimising the energy consumption and use of other resources during the lifecycle of high-risk AI systems, and by promoting energy-efficient development of GPAI models (with related standards to be integrated in the coming years). In addition, Art. 95 encourages to implement environmental sustainability, including energy-efficient programming and techniques for efficient design, training and use of AI. Providers of GPAI models must also be required to include in their technical documentation a detailed description of the computational resources used to train their models, as well as their energy consumption, as outlined in Annex XI of the AI Act.

More generally, recital 27 connects environmental well-being to the AI High-Level Expert Group Ethics Guidelines (*infra* 10.7). It emphasises that AI systems must be developed and used in a sustainable and environmentally friendly manner, ensuring they benefit all human beings. This principle extends beyond energy consumption to focus on broader human-centric, trustworthy AI design.

For ACHILLES, sustainability is a key objective, and various measures are being considered to **reduce energy consumption** during the design, deployment and use of AI systems, such as reducing time to train models (e.g., improved hyperparameter tuning and more efficient datasets), or making inference more energy-efficient (e.g., model pruning) while preserving the performance.

4.1.7 Union harmonisation legislation

Art. 8(2) AI Act specifies that in addition to the obligations set forth in the AI Act, requirements under Union Harmonisation legislation still apply. This is particularly important for the healthcare use case, which will likely fall under Union harmonisation legislation in Section A of Annex I (including the Medical Devices Regulation).

[56] Gils, T., 2024, 'A Detailed Analysis of Article 50 of the EU's Artificial Intelligence Act', *The EU Artificial Intelligence (AI) Act: A Commentary*, Kluwer.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

4.2 International regulatory initiatives

Beyond the EU AI Act, it is important to briefly review a few relevant international regulatory initiatives as they may offer useful insights that complement the AI Act. However, the EU AI Act remains the primary source to consider, given its direct applicability in EU Member States and its cross-border effects.

4.2.1 UNESCO Recommendations on the Ethics of AI

In 2021, UNESCO issued its Recommendations on the Ethics of AI.[57] Although not legally binding, these recommendations provide useful – and even normative – recommendations about AI technologies. They urge governments to establish the necessary institutional and legal frameworks to govern AI technologies and ensure they contribute to the public good. The recommendations promote human rights, human dignity, and environmental sustainability, and advance principles of transparency, accountability, and rule of law online. They act as guidelines to all governments when drafting AI laws and strategies.

4.2.2 OECD Recommendations on AI

In 2019, the OECD adopted its Recommendations on AI, updated in 2024.[58] These principles serve as an intergovernmental standard on AI, though they are not legally binding. They promote innovative, trustworthy AI that respects human rights and democratic values. The principles include: Inclusive growth, sustainable development and well-being; Human rights and democratic values, including fairness and privacy; Transparency and explainability; Robustness, security and safety; Accountability.

The OECD also provides recommendations for policymakers, such as: Investing in AI research and development; Fostering an inclusive AI-enabling ecosystem; Shaping an enabling interoperable governance and policy environment for AI; Building human capacity and preparing for labour market transformation; International co-operation for trustworthy AI.

4.2.3 Council of Europe Framework Convention on AI

At the supranational level, the Council of Europe adopted its Framework Convention on AI and Human Rights, Democracy and the Rule of Law in 2024.[59] It is the first-ever international legally binding treaty, focusing on human rights, democracy and the rule of law. It ensures these principles are respected throughout the AI lifecycle. The AI Convention includes fundamental principles, such as Human dignity and individual autonomy; Equality and non-discrimination; Respect for privacy and personal data protection; Transparency and oversight; Accountability and responsibility; Reliability; Safe innovation.

[57] UNESCO, 2022, *Recommendation on the Ethics of Artificial Intelligence*, https://unesdoc.unesco.org/ark:/48223/pf0000381137.
[58] OECD, 2019, *AI Principles*, https://www.oecd.org/en/topics/ai-principles.html.
[59] Council of Europe, 2024, *The Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, https://rm.coe.int/1680afae3c.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

It also includes an obligation to conduct risk and impact assessments in respect of impacts on human rights, democracy and rule of law, and the need to establish measures to mitigate risks.

As for the scope, the Convention applies to the use of AI systems by public authorities, including private actors acting on their behalf, as well as private actors. Member states to the Convention have two options regarding the private sector: they can either directly apply the relevant provisions to them or take other measures to comply.

The Convention is open for accession by the 46 Council of Europe Member States, the EU and other states that are not members of the Council of Europe. It has opened for signature on 5 September 2024. Once five states have given their consent to be bound by the Convention (e.g. after ratification), it will enter into force.

4.3 Standards

In addition to European, international and supranational initiatives, standards play an important role in regulating the design, development, deployment and use of AI systems. Normally, standards are technical and focus on the implementation of specific technical aspects (and should not extend to interpreting legal requirements).[60] Standards can enhance transparency, data quality, reliability, and so on, depending on their exact scope.

The International Organisation for Standardisation (ISO) is one of the most prominent non-governmental standardisation organisations, with experts drafting standards. It operates at the international level. ISO's standards are not legally binding but can help with compliance by establishing a presumption of conformity. Together with the Electrotechnical Commission (IEC), ISO sets standards for AI safety, transparency and risk management, among other areas. Nevertheless, a major issue with ISO standards is that they are not free.

For the EU AI Act, CEN-CENELEC – a joint committee of two standardisation organisations in the EU – is currently drafting standards to help interpret and implement some of the Act's requirements. While compliance with these standards is not mandatory, it creates a legal presumption that organisations conforming to them meet the related requirements in the AI Act (Art. 40 AI Act). CEN-CENELEC is currently working on standards in various areas, including terminology for AI, AI risk management, FRIA, natural language processing, AI ethics and social concerns, environmental impact of AI, transparency, system logging, nudging, AI conformity assessment and so on.[61] It is expected that CEN-CENELEC will

[60] BEUC, *Regulating AI To Protect The Consumer – Position Paper on the AI Act*, https://www.beuc.eu/sites/default/files/publications/beuc-x-2021-088_regulating_ai_to_protect_the_consumer.pdf.
[61] For a full overview of work programme of CEN-CENELEC, please see: CEN-CENELEC, 2022, *About CEN*, https://standards.cencenelec.eu/dyn/www/f?p=205:22:0::::FSP_ORG_ID,FSP_LANG_ID:2916257,25&cs=1827B89DA69577BF3631EE2B6070F207D.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

build on the ISO/IEC AI standards and ensure that their standards align with EU fundamental rights and specific requirements, and tailor them to meet the legal requirements of the EU AI Act.

Below, two recently adopted ISO AI standards are outlined.

4.3.1 ISO/IEC 42001: AI Management System

ISO/IEC 42001: AI Management System is the first international standard to establish a framework for implementing, maintaining, and improving an Artificial Intelligence Management System (AIMS) within organisations.[62] The standard provides an integrated approach to **managing AI projects**, which is essential as AI technology continues to evolve rapidly. It offers a comprehensive approach to AI risk assessment and mitigation.

ISO/IEC 42001 is designed to help businesses build trust and credibility by ensuring that AI is used safely and responsibly. The standard offers guidance for organisations to address challenges related to ethics, transparency and continuous learning. It does not focus on specific AI applications but rather on managing AI-related risks and opportunities across the organisation.

Although it is behind a paywall, ISO/IEC 42001 includes several requirements for organisations. These include maintaining high standards for continual improvement and monitoring the performance of AIMS. Organisations are also required to implement processes to identify, analyse, evaluate, and monitor risks throughout the entire lifecycle of the system. This includes assessing the consequences and risks associated with the AI system and implementing measures to improve performance. However, for more specific aspects of AI management, such as AI model validation, it is better to check more specific standards to ensure compliance with individual components of robust AI systems.[63]

4.3.2 ISO/IEC TR 2408:2020 Information Technology – AI – Overview of trustworthiness in artificial intelligence

Another ISO standard, ISO/IEC TR 2408:2020 Information Technology – AI – Overview of trustworthiness in artificial intelligence [64] analyses and provides guidelines on factors that can impact the trustworthiness of AI systems. The standard surveys existing approaches to **improve trustworthiness** in technical systems and AI applications, offering mitigation measures to improve the overall reliability of AI systems, such as the implementation of robustness testing and fault tolerance mechanism, and

[62] ISO/IEC, 2023, *ISO/IEC 42001:2023 - AI management systems*, https://www.iso.org/standard/81230.html; ISO, 2025, *AI Risk Assessments Under ISO/IEC 42001: A Practical Guide*, https://iso-docs.com/blogs/iso-42001-artificial-intelligence-management-system-aims/ai-risk-assessments-under-iso-iec-42001-a-practical-guide.
[63] KPMG, 2025, ISO/*IEC 42001: The latest AI management system standard*, https://kpmg.com/ch/en/insights/artificial-intelligence/iso-iec-42001.html.
[64] ISO/IEC, 2020, *ISO/IEC TR 24028:2020 Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence*, https://www.iso.org/standard/77608.html.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

the use of data quality assurance processes to minimise errors or biases during training and operation.[65]

[65] ISO, 2025, *Towards a trustworthy AI*, https://www.iso.org/news/ref2530.html#:~:text=To%20address%20issues%20of%20trust%20in%20artificial%20intelligence,concerns%20related%20to%20trustworthiness%20and%20provides%20practical%20solutions.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

5 PRIVACY AND DATA PROTECTION

5.1 General Data Protection Regulation

5.1.1 Introduction

The development of the ACHILLES IDE involves various activities that require the processing of personal data. For instance, in WP7, human participation will be part of research activities in the use case validation process. In this context, personal data, such as name, surname, email, and others, will be collected as part of training datasets. Certain use cases, such as the healthcare and identity verification ones, require the processing of specific categories of personal data, namely health and biometric data respectively. In addition, AI systems developed through the IDE must comply with data protection regulation.

Besides the GDPR, private and family life and protection of personal data are also protected by Convention 108+ of the Council of Europe.[66] It is the only legally binding international instrument in the data protection field. It applies to all data processing activities by both private and public sectors, and lays down principles for processing and data subject rights. Although it is not subject to the judicial supervision of ECtHR, its case law is taken into consideration. All EU Member States have ratified Convention 108+.[67] Since the GDPR's principles closely align with the Convention's principles – sometimes even expanding them, and since the GDPR has a direct effect in the EU, this analysis will focus on the GDPR.

The GDPR operationalises fundamental rights related to privacy and personal data protection, and provides a clear set of concrete requirements and obligations. The following part will analyse the GDPR's scope, principles and rights, as well as the role of key actors, DPIAs and other relevant aspects for the ACHILLES project.

Importantly, ETICAS will provide both more ethical and technical support for the GDPR requirements. ETICAS will also appoint an ethical advisory board to oversee the design and implementation of measures addressing bias mitigation, and conduct a technical analysis of GDPR compliance within the project. In addition, with regard to research data, the first version of the Data Management Plan (DMP) is currently being drafted, due in M6, and will include information about data licensing and availability, re-use of data, duration of data for re-use, reproducibility of research output and so on.

5.1.2 Scope

[66] Council of Europe, 2018, *Convention 108 + Convention for the protection of individuals with regard to the processing of personal data*, https://rm.coe.int/convention-108-convention-for-the-protection-of-individuals-with-regar/16808b36f1.

[67] FRA, 2018, *Handbook on European data protection law*, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

The **territorial scope** of the GDPR is broad, meaning that the ACHILLES project activities would most likely fall under it (Art. 3 GDPR). The GDPR applies to every processing of personal data if (1) the data controller(s) or processor(s) are established in the EU, regardless where processing takes place, (2) when data controller(s) or processor(s) are not established in the EU, but the processing activities are related to the offering of goods or services, or the monitoring of their behaviour that is taking place in the EU, or (3) when public international law makes GDPR applicable.

In terms of **material scope**, the GDPR applies when two cumulative conditions are met, namely (1) the data is personal data, and (2) the personal data is processed.[68] Both notions of "personal data" and "processing" are broadly interpreted.

Personal data is defined as "*any information relating to an identified or identifiable natural person, the data subject*" (Art. 4(1) GDPR). Any information entails all possible information directly or indirectly linked to a data subject, irrespective of its nature, as long as the information is *about* the individual. This could include, for instance, name, telephone number, address, social security number, vehicle registration number and so on. Importantly, actual identification of data subjects is not required to be considered personal data. Instead, it is sufficient that a person is identifiable, meaning that the available data, when combined with other reasonable means, could allow for their identification. The assessment of identifiability thus depends on the context.

Pseudonymised data fall within the scope of the GDPR. Art. 4(5) GDPR defines "pseudonymisation" as "*the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person*". They are considered personal data, as they still allow re-identification of the individual as additional information exists. Whether the data recipient has the reasonable means to re-identify an individual depends on the context and concrete additional information.[69] Pseudonymisation is a privacy-enhancing measure that helps comply with the data minimisation principle, data protection by design and security obligations under the GDPR.[70]

---

[68] Voigt, P. & von dem Bussche, A., 2017, T*he EU General Data Protection Regulation*, Springer, pp. 9-30.
[69] Recital 26 GDPR; Finck M. & F. Pallas, F., 2020, 'They who must not be identified – distinguishing personal from non-personal data under the GDPR', *International Data Privacy Law* 2020, pp. 11-36; CJEU C-582/14, Breyer, 19 October 2016, ECLI:EU:C:2016:779; CJEU T-557/20, SRB v EDPS, 16 April 2023, ECLI:EU:T:2023:219.
The judgement is now under appeal: https://curia.europa.eu/juris/document/document.jsf?text=&docid=276483&pageIndex=0&doclang=EN&mode=req&dir=&occ=first&part=1&cid=323544.
.
[70] EDPB, 2025, *Guidelines 01/2025 on Pseudonymisation*, https://www.edpb.europa.eu/system/files/2025-01/edpb_guidelines_202501_pseudonymisation_en.pdf.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

In contrast, the GDPR does not apply to anonymous data. Anonymous information refers to information that does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable.[71] The threshold to be considered anonymous is high: it must be impossible to retrace the identity of the data subject, and the anonymisation process must be irreversible. In practice, this will require a case-by-case analysis.[72]

Second, "processing" is defined as "*any operation or set of operations which is performed on personal data or on sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction*" (Art. 4(2) GDPR). This broad definition covers any action involving personal data, from collection and storage, to modification, sharing or even deletion. Essentially, whenever an organisation interacts with personal data in any way, it is engaged in processing, making it subject to the GDPR obligations.

5.1.3 Data subjects, data controllers and data processors

Under the GDPR, **data subjects** are the beneficiaries of data protection rules. They are exclusively natural persons.[73] Data subjects are granted extensive rights to protect their privacy and ensure control of their personal data.

Data controllers and **data processors** are the two other main actors under the GDPR. Given the processing activities during the ACHILLES project, it will be important to consider who is responsible for ensuring compliance with the GDPR in which particular activities. For each activity involving the processing of personal data, a data controller and, possibly, a data processor should be designated. The data controller is the main addressee of the GDPR and responsible for complying with almost all principles and rules under the GDPR. It refers to the natural or legal person who determines the purposes (the *why*) and the means (the *how*) of the processing (Art. 4(7)). Essential means could include decisions on what data will be processed, how long the data will be retained, who the recipients of the data are, and which categories of data subjects are involved. In turn, data processors do not determine the purpose of the processing but carry out processing activities *on behalf of* the controller (Art. 4(8)). To qualify as a data processor, two conditions must be met: (1) the processor is a *separate entity* in relation to the controller, and (2) the processor processes personal data *on behalf of* the controller. A separate entity implies that the processor is an external organisation, e.g. not an employee or under the direct authority of the data controller. The processing must take place on behalf of the controller, meaning that the processor implements the controller's instructions and does not process data for its

[71] Recital 26 GDPR; CJEU T-557/20, Single Resolution Board v. EDPS, 26 April 2023, ECLI:EU:T:2023:219.
[72] Stalla-Bourdillon S. & Knight, A., 2017, 'Anonymous Data v. Personal Data — A False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data', *Wisconsin International Law Journal*.
[73] FRA, 2018, *Handbook on European data protection law*, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG.

own purpose(s). In case of a data processor, the relationship between the data controller and data processor must be governed by a **contract** or another legal act in a written manner, in accordance with Art. 28(2). *Non-essential* aspects of the processing means can be delegated to the processor, such as more practical aspects of implementation, or detailed security measures.[74]

In some cases, multiple controllers may be involved in the same processing activity, forming **joint controllers**. According to Art. 26 GDPR, joint controllers exist when two or more entities jointly determine both the purpose and means of processing. The key criterion for joint control is the shared decision-making, which is to be assessed on factual circumstances. The CJEU has interpreted this in a broad manner, considering any actor that has made the processing possible by contributing to it.[75] To ensure transparency and accountability, joint controllers should clearly define their roles and responsibilities. [76] The European Data Protection Board (EDPB) recommends drafting a binding document (such as a contract or another legally binding act) that each party's responsibilities, clarifies obligations towards each other and towards data subjects and ensures compliance with GDPR principles.[77]

In the age of automation and AI, determining who is responsible for processing activities can be complex. Due to the broad interpretation of controller and joint controller, AI systems often involve multiple actors across different stages of the processing chain. This complexity raises the risk that almost everyone could be classified as a joint controller. To address this, some experts have suggested to break down the complex processing operations into smaller distinct processing activities before assessing the roles.[78] For instance, a company that develops and commercialises ML models can determine why certain training datasets are processed (e.g. to monetise the model) and essential and non-essential means (e.g. what personal data is included and how the training process is implemented). In that case, this company is the sole data controller for training dataset processing. If another company provides training datasets in exchange for access to the trained model, both companies *jointly* decide on the processing and can be classified as joint controllers. Imagine if a company uses the AI model for a specific purpose (e.g. risk assessments or medical diagnostics), then that company determines why and how the AI model processes new data, making it the sole processor

[74] Foulsman, M., Hitchen, B. & Denley, A., 2019, *GDPR. How to Achieve and Maintain Compliance*, Routledge, pp. 21-23.
[75] Dewitte, P., 2024, 'AI Meets the GDPR, Navigating the Impact of Data Protection on AI Systems', *The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence*, pp. 133-157; CJEU C-131/12 Google Spain, 13 May 2024, ECLI:EU:C:2014:317.
[76] Cimina, V., 2021, 'The data protection concepts of 'controller', 'processor' and 'joint controllership' under Regulation (EU) 2018/1725', *ERA Forum*, pp. 639-654.
[77] EDPB, 2020, *Guidelines 07/2020 on the concepts of controller and processor in the GDPR, version 1.0*, https://edpb.europa.eu/sites/default/files/consultation/edpb_guidelines_202007_controllerprocessor_en.pdf.
[78] Dewitte, P., 2024, 'AI Meets the GDPR, Navigating the Impact of Data Protection on AI Systems', *The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence*, pp. 133-157.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

for that activity. However, the original model provider may still exert influence over how the algorithm functions.

5.1.4 Principles and Rights

Whenever personal data is processed during the ACHILLES project, the GDPR requires responsible partner(s) to comply with the six principles as enshrined in Art. 5 GDPR.[79]

Lawfulness, fairness and transparency	Lawfulness requires partners to determine an appropriate *legal basis* before processing personal data in accordance with Art. 6 GDPR. For each project activity where the processing of personal data takes place, a legal basis should be chosen before the processing activity takes place. For ACHILLES, the appropriate legal basis will likely be consent. Fairness requires data controllers to assess data subjects' interests and meet reasonable expectations regarding the processing activity. Specifically with regard to biases and AI training, fairness requires controllers to create or use datasets with training data that are adequate and have a fair representation of the real world. Therefore, they should implement accountability and oversight mechanisms, such as audits, and keep supporting documentation.[80] Transparency requires data controllers to inform data subjects about the processing activities and their rights, as per Art. 12-24 GDPR. Information should be communicated to the data subjects about the collection, use and storage of their data in a concise, transparent, intelligible and easily accessible form. Transparency brings about awareness of what is being done with personal data and enables individuals to exercise their rights.[81]
Purpose limitation	Data controllers must specify the purposes for which personal data are processed for each processing activity before communicating it to the data subjects (recital 39 GDPR). The data must be collected for a *specific, explicit, and legitimate purpose, which is* determined at

[79] Voigt P. & von dem Bussche, A., 2017, T*he EU General Data Protection Regulation*, Springer, pp. 87-92.
[80] EDPS, 2024, *Generative AI and the EUDPR. First EDPS Orientations for ensuring data protection compliance when using Generative AI systems*, https://www.edps.europa.eu/system/files/2024-06/24-06-03_genai_orientations_en.pdf.
[81] FRA, 2020, *Your rights matter: data protection and privacy*, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-fundamental-rights-survey-data-protection-privacy_en.pdf.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	the time of the collection of the personal data. This means that any processing for undefined purposes is unlawful. Further processing of personal data that is incompatible with the original purpose is unlawful. Only further processing within project's targets is allowed (*infra* 5.1.6). Following the principle of purpose limitation, project partners should take appropriate measures to ensure that only personal data relevant for the project objectives are collected and processed. The responsible project partners must specify and justify the purpose for each different processing activity.
Data minimisation	Data controllers must limit the amount of data collected and processed to what is strictly necessary to achieve the specified purposes and should not constitute a disproportional interference with the interest, rights and freedoms of the data subjects. The period for which the personal data is stored must be limited to the minimum. However, in the age of AI and big data, big data's business model may often contradict the principle of data minimisation, as AI and big data activities require more and more data, often for unspecified purposes.[82] Therefore, quality over quantity should be emphasised: structured datasets, supervised learning processes and regular monitoring.[83] Following the data minimisation principle, it should be examined how to achieve the purpose with the least amount of personal data. Partners could, for instance, also consider using anonymisation techniques.
Accuracy	Accuracy requires data controllers to regularly check and rectify or erase inaccurate data. Partners should commit themselves to take reasonable steps to ensure that the data processed is accurate and kept up to date. For generative AI systems, accuracy is particularly challenging as data are often sourced from multiple channels without direct

[82] FRA, 2018, *Handbook on European data protection law*, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG.
[83] EDPS, 2024, *Generative AI and the EUDPR. First EDPS Orientations for ensuring data protection compliance when using Generative AI systems*, https://www.edps.europa.eu/system/files/2024-06/24-06-03_genai_orientations_en.pdf.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	validation. It is therefore important to ensure that content and structure datasets used for training are and remain accurate. Partners could engage in regular monitoring, human oversight and keeping documentation, or use validation sets during training to see how the system will perform.[84]
Storage limitation	Storage limitation requires data controllers to restrict the periods of time to store personal data to no longer than is necessary for the purposes. Once the personal data is no longer necessary for the processing purposes, they must be deleted or anonymised/pseudonymised.
Integrity and confidentiality	Integrity and confidentiality require that data is processed to ensure appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures. According to this principle, partners should delete the personal data once it has been used for its intended purpose. Partners may also implement security measures to prevent unauthorised access to personal data, for example by developing security policies on identification, authentications and authorisation.

Table 1 – GDPR Principles

Art. 5(2) GDPR establishes that the controller is responsible for ensuring compliance with the six principles and must be able to demonstrate this compliance. To satisfy this accountability principle, appropriate documentation is required.

In addition to accountability, data controllers must implement data protection **by design and by default**.[85] This means that data protection principles must already be embedded into the design of any activity from the outset. In the context of developing the AI tool, this means that, even in the early design phase, careful consideration must be given to the types of personal data the tool may collect from users, the application of GDPR principles, and the security measures required to protect personal data. Data protection by design obliges data controllers to implement both technical and organisational measures, such as pseudonymisation and data minimisation, at every stage, from determining the means of processing to the actual processing itself. The data controller must also ensure that, by

[84] FRA, 2018, *Handbook on European data protection law*, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG.
[85] Lievens E. & van der Hof, S., 2018, 'The importance of privacy by design and data protection impact assessments in strengthening protection of children's personal data under the GDPR', *Communications Law*, pp. 33-43.

default, only the personal data necessary for a specific processing activity is processed. This includes limiting the scope of data collection, restricting storage duration and ensuring that access to data remains strictly controlled.

Closely related is the **principle of security** of processing, as outlined in Art. 32 GDPR. This article requires both controllers and processors to implement appropriate technical and organisational security measures to the risks associated with processing activities. These could include pseudonymisation and encryption of personal data; ensuring permanent confidentiality, integrity, availability and resilience of the processing systems and services; the ability to restore access to personal data in case of a personal or technical incident; or a process for regularly testing, assessing and evaluating the effectiveness of the measures. The appropriateness of the measures depends on the risks of the processing.

Specifically to mitigate the risk of **discrimination**, it is crucial to address potential biases from the outset rather than attempting to correct them retroactively. Simply excluding information related to protected characteristics, such as gender, is not a sufficient safeguard against discrimination, as other seemingly neural data points, such as postal codes, can serve as proxies for sensitive attributes (*supra* 3.5). This reinforces the need for a robust approach to data protection by design and by default. Merely omitting certain data does not eliminate bias. Instead, controllers must adopt appropriate technical and organisational measures to integrate safeguards that proactively prevent discrimination. For instance, fairness and non-discrimination principles should already be embedded during the design phase into the data processing framework.[86]

The GDPR grants different rights to data subjects. Below follows a general overview of the rights.

Art. 12 – **Transparent** information, communication and modalities for the exercise of the rights of the data subject	Data controllers must provide any information relating to the processing to the data subject in a concise, transparent, intelligible and easily accessible form, using clear and plain language. This includes not only the purposes and legal basis for processing, but also their rights. The features of data processing systems must make it possible for data subjects to really understand what is happening with their data. It includes informing data subjects on the details of the identity and contact information of the data controller, data retention periods, recipients or categories of recipients of the data, and whether the

[86] FRA, 2018, *#BigData: Discrimination in data-supported decision making*, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-focus-big-data_en.pdf.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	data will be transferred outside the EU or to an international organisation.[87] Even in cases of complex data processing, the complexity of the processing should not, in itself, preclude the data controller from providing data subject with clear explanations on the objectives and analytics used in the data processing.[88]
Art. 13 – **Information** to be provided where personal data are collected from the data subject	To ensure fair and transparent processing, the GDPR requires that data controllers provide the data subject with meaningful information about the logic involved in automated decision-making, including profiling (Art. 13(2)(f)). The data controller, when data is collected from the data subject, must provide information on the controller's identity, the contact details of the data protection officer (when applicable), the processing purposes and legal basis, the recipients of the personal data (if any) and on data transfers (when applicable). The data controller must provide further information necessary to ensure fair and transparent processing, namely the identity and contact details of the controller, the contact details of the data protection officer, the purposes of the processing, the recipients or categories of the personal data, the storage period, the existence of the right to request from the controller access to and rectification or erasure of personal data or restriction of processing, the right to data portability, the right to withdraw consent when the legal basis is consent, the right to lodge a complaint with a supervisory authority, the existence of automated decision-making, including profiling.
Art. 14 – **Information** to be provided where personal data have not been obtained from the data subject	Where personal data have not been obtained directly from the data subject, the controller must provide similar information as under Art. 13, together with the source from which the personal data originate. The information must be provided within a reasonable time after obtaining the personal data and, at the latest, within one month.

[87] FRA, 2020, *Your rights matter: data protection and privacy*, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-fundamental-rights-survey-data-protection-privacy_en.pdf.
[88] Article 29 Data Protection Working Party, 2018, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, https://ec.europa.eu/newsroom/article29/items/612053.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Art. 15 – Right of **access** by the data subject	Data subjects have the right to know whether their personal data is being processed and to access the personal data, as well as information on the processing purposes, categories of personal data, recipients, storage period, sources. They may also request a copy of the personal data undergoing processing, in intelligible form. The right to access is important because it allows individuals to understand how their data is being used and to exercise further rights, such as rectification or erasure.[89]
Art. 16 – Right to **rectification**	Data subjects have the right to obtain the rectification of inaccurate or incomplete personal data from the data controller.
Art. 17 – Right to **erasure** ('right to be forgotten')	Data subjects have the right to obtain the erasure of personal data from the data controller if one of the following grounds apply: ▪ The personal data is no longer necessary for the processing purpose ▪ The data subject withdraws his or her consent in case the processing activity relies on the legal basis of consent ▪ The personal data have been unlawfully processed The right to erasure reinforces the principle of purpose limitation by ensuring that data is not retained beyond necessity.
Art. 18 – Right to **restriction** of processing	Data subjects have the right to limit or restrict how their personal data is used where one of the following applies: ▪ The data subject contests the accuracy of the personal data ▪ The processing is unlawful, and the data subjects oppose the erasure of their personal data and request the restriction of their use ▪ The data controller no longer needs the personal data for the processing purpose, but the data subjects request them in relation to a defence or legal claim

[89] FRA, 2020, *Your rights matter: data protection and privacy*, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-fundamental-rights-survey-data-protection-privacy_en.pdf.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	▪ The data subject has objected the processing on the basis of legitimate interests of the data controller. This requires the ability of the controller(s) and/or processor(s) to temporarily transfer the personal data to another system or make them unavailable.
Art. 19 – **Notification obligation** regarding rectification or erasure of personal data or restriction of processing	Data controllers must notify data subjects of any rectification or erasure of personal data or restriction of processing, unless this proves impossible or involves disproportionate efforts.
Art. 20 – Right to **data portability**	Data subjects have the right to obtain the personal data from the data controller and transmit this data to another data controller where the processing is based on the legal basis of consent and is carried out through automated means.
Art. 21 – Right to **object**	Data subjects have the right to object to the processing of their personal data when the processing is based on the legal basis of public interest or legitimate interests, or direct marketing purposes.[90]
Art. 22 – **Automated individual decision-making**, including profiling	Data subjects have the right not to be subject to automated decisions without human involvement. It entails a general prohibition on fully automated decision-making. This concerns decisions *solely* based on automated processing, which produces legal effects concerning the data subject. Data controllers may be exempted from such prohibition only in three specific cases: when the decision is: 1) necessary for the performance of a contract between the data subject and the controller, 2) permitted by an EU or national law, or 3) based on explicit consent.[91]

Table 2 – GDPR Data Subjects' Rights

In the context of ACHILLES, the **lawfulness principle** (Art. 5(1)(a) GDPR) requires partners to choose a valid legal basis. For the use case validation WP, **consent** has been identified as the most appropriate legal basis. Consent also appears to be the most appropriate legal basis for using the ACHILLES IDE, especially when AI developers choose to use this environment. Art. 4(11) GDPR defines consent in a

[90] FRA, 2018, *Handbook on European data protection law*, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG.
[91] FRA, 2018, *Handbook on European data protection law*, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG.

strict manner, as "*any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her*".[92] Art. 7, along with recitals 32, 33, 42 and 43, further clarify the conditions of consent:[93]

- Consent should given by a *clear affirmative act*.
- Consent must be *freely* given. This means that data subjects have a real choice and control. Data subjects should not feel compelled to consent, e.g. due to an imbalance of power.
- Consent must be *specific*. This implies that the consent must be given for a specific purpose. If there are multiple purposes, the consent must be given for all of them.
- Consent must be *informed*. This means that the request for consent and the explanation of the data processing activities and their purposes are described in clear and plain language without technical jargon. Data subjects should also be aware of the controller's identity and the type of data collected and used.
- Consent must be an *unambiguous indication* of the data subject's agreement. This means no doubt may exist about whether the data subject consented. An unambiguous indication of agreement can be made through a written or an oral statement. Nevertheless, for the controller to be able to demonstrate consent, a written statement is preferable. Silence, pre-ticked boxes or inactivity do not constitute valid consent. If the consent is given by electronic means, the request must be clear, concise and not unnecessarily disruptive to the use of the service for which it is provided.
- Consent can be *withdrawn* at any moment. The withdrawal of the consent must be as easy as obtaining the consent.

As for the human participants involved in the WP7 activities (use cases validation), only adult participants who have provided their consent (to ensure autonomy) will be included, and informed consent forms will be used. No vulnerable people are recruited. As outlined in the ACHILLES project proposal, the following **procedure** will be followed to obtain valid and voluntary consent:

1) Interested persons will be *informed* about the project and get details about their involvement, rights and purpose of the collection of personal data. The voluntary nature of their participation and the confidentiality of the information will be emphasised.
2) Participants are given an opportunity to *ask questions*.
3) A *cooling-off* period of at least seven days will be provided to allow participants to consider their decision.
4) Participants may *ask questions* again.
5) Participants provide their *informed consent*.

[92] EDPB, 2020, *Guidelines 05/2020 on consent under Regulation 2016/679, Version 1.1*, https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf; Breen, S., e.a., 2020, 'GDPR: Is your consent valid?', *Business Information Review*, pp. 19-24.
[93] See also D4.4 from ETICAS clarifying some of the GDPR principles in more detail.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

As for the **data minimisation principle**, technical and operation measures must be implemented to ensure compliance. This includes clearly defined purpose(s), limited access and storage to personal data, pseudonymisation and other (more technical) measures. For example, access to personal data obtained must be limited to only the necessary personnel, and the personal data must be kept secure, password-protected and not disclosed to anyone outside the consortium. Consent forms must be stored in a locked filing cabinet in an office with limited access, or online in a secure environment. Notes taken must be pseudonymised to further protect participants' identities.

5.1.5 Special categories of personal data

The GDPR provides additional rules for the processing of special categories of personal data, which inherently carry higher risks to data subjects when processed, given their sensitive nature. These data concern racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic and biometric data for the purpose of identifying a natural person, data concerning health, a person's sex life or sexual orientation. The processing of these types of data is, **in principle, prohibited** under Art. 9(1) GDPR, and can only occur under a limited number of conditions.

This is particularly relevant for ACHILLES in the context of the use case validation WP. For instance, in the healthcare and identity verification use cases, sensitive health data[94] and biometric data,[95] respectively, are involved. In addition, other AI systems later developed through the ACHILLES IDE may involve the processing of sensitive personal data.

The GDPR foresees some **exceptions** to the general prohibition on processing sensitive data, such as **explicit consent** from data subjects. Art. 9(2)(h) provides a specific legal basis for processing sensitive medical data without consent, i.e. when the processing of medical data is required for the purpose of preventive medicine, medical diagnosis, provision of care or treatment, or management of healthcare services. However, this exception is only applicable when processing is carried out by a healthcare professional subject to obligations of professional secrecy – which is not the case in ACHILLES. Therefore, explicit consent will be required in these scenarios.

Explicit consent is required when sensitive data is involved, given the higher risks and, thus, the greater level of control needed over the data. In addition to all requirements for the 'regular' consent as mentioned hereabove, explicit consent entails that data subjects must give an **express statement** of consent, most straightforwardly by confirming consent in written statement, and ideally also signed by the data subject. Oral means or electronic means can also work for obtaining explicit consent, but can make it more difficult for the controller to prove the conditions for valid consent were fulfilled. Another

[94] Art. 3(15) defines 'data concerning health' as personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status.
[95] Art. 3(14) defines 'biometric data' as personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

potential solution is implementing a two-stage verification process to ensure consent is properly obtained and documented.[96]

5.1.6 Further processing

Art. 6(4) GDPR clarifies that further processing of personal data is, **in general, prohibited when it is incompatible with the original purpose(s)** for which the data was processed. This is closely tied to the purpose limitation principle in Art. 5(1)(b). Specifically, data controller(s) must clearly define the purpose(s) for which they intend to process personal data. They cannot process personal data for different purposes via the same processing operation(s). This provision is relevant for the ACHILLES use cases, in particular when processing personal data that is already available (either publicly or in datasets held by other organisations), and has been collected for a specific purpose.

However, there are two **exceptions** to this general prohibiting on further processing. The first is when further processing is based on the **consent** of the data subjects. If the original consent does not cover further processing, the data controller must find another legal basis. The second exception arises when processing is **required under Union or Member State law**, as a necessary and proportionate measure in a democratic society. In this case, further processing is allowed, irrespective of compatibility with the initial purpose. In addition, compatibility is presumed when data are to be processed for "*archiving purposes in the public interest, scientific or historical research purposes or statistical purposes*" (Art. 89). However, even in these cases, the data controller must still ensure there is a valid legal basis for the processing operation as outlined in Art. 6 GDPR.

If these exceptions do not apply, the data controller is required to conduct a **compatibility assessment** for the further processing activity, as introduced in art. 6(4) (see also recital 50).[97] In this assessment, the controller can proceed with further processing when it is ***necessary and proportionate*** for another purpose that is compatible with the purpose for which the personal data were initially collected, taken into account, among others: (a) any link between purposes for which personal data have been collected and purposes of intended further processing (i.e. whether they are closely aligned); (b) context in which personal data have been collected; (c) nature of personal data (especially when special categories of personal data concerned); (d) possible consequences of intended further processing for data subjects; (e) existence of appropriate safeguards, which may include encryption or pseudonymisation.

---

[96] EDPB, 2020, *Guidelines 05/2020 on consent under Regulation 2016/679*, https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf.

[97] Becker, R., e.a., 2022, 'Secondary Use of Personal Health Data: When Is It "Further Processing" Under the GDPR, and What Are the Implications for Data Controllers?', *European Journal of Health Law*; European Parliament, 2020, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, https://op.europa.eu/en/publication-detail/-/publication/dc544697-19b8-11ec-b4fe-01aa75ed71a1/language-en; FRA, 2018, *Handbook on European data protection law*, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

5.1.7 Synthetic data

As synthetic data, i.e. artificial data generated from original data to train AI model to reproduce characteristics and structure of original data – is also planned to be used in the training of AI systems within the ACHILLES project, a key question arises regarding its status under the GDPR. Synthetic data offers the advantage to deliver results closely **mimicking real data**. From a data protection perspective, synthetic data is seen as a privacy-enhancing measure because it does not disclose personal data and can help mitigate biases, thus promoting fairness.[98]

However, the risk exists that individuals could potentially be identified through synthetic data, especially if it retains enough structural similarity to the original data. Whether or not synthetic data could be considered pseudonymised or anonymised data depends on the circumstances. Some have argued that if correctly generated, synthetic data has no one-to-one mapping back to the real individuals and can thus be considered anonymous because it is indistinguishable from the original data. In contrast, others have argued that it should not be considered anonymous data because one-to-one relationships are still possible to derive when a synthetic data set still has characteristics of original data with high accuracy.[99] It has been argued that if the source data is personal data, then the output data (i.e. synthetic data) remains personal data unless it can be demonstrated that re-identification threats are minimal.[100] In conclusion, the classification of synthetic data will depend on **how much it deviates from the original data** and whether it can maintain anonymity over time. In any case, privacy assurance measures have to be taken to ensure that synthetic data does not inadvertently reveal actual personal data.

5.1.8 Data Protection Officer and Data Protection Impact Assessment

The final section will discuss two concepts that are relevant in light of privacy and data protection within the ACHILLES project: the role of the Data Protection Officer (DPO) and the use of a Data Protection Impact Assessment (DPIA). For more details on technical and practical implementation, please see **D4.4**.

Under certain conditions, a **DPO** must be appointed by the data controller or processor. The DPO is a person legally obliged to oversee the company's data activities. This person must have a high level of expert knowledge of legislation, practices and GDPR compliance. Some of the duties include continually tracking and monitoring GDPR compliance for the organisation, implementing training for employees about compliance and performing GDPR audits; implementing and performing data

[98] Riemann, R., 2025, *Synthetic Data*, https://www.edps.europa.eu/press-publications/publications/techsonar/synthetic-data_en.
[99] Stadler, T., Oprisanu, B. & Troncoso, C., 'Synthetic Data -- Anonymisation Groundhog Day', arxiv.2011.07018.
[100] Global Alliance, 2024, *GDPR Brief: when are synthetic health data personal data?*, https://www.ga4gh.org/news_item/when-are-synthetic-health-data-personal-data/#:~:text=The%20key%20question%20is%20whether%20synthetic%20data%20fall,law%20as%20%E2%80%98personal%20data%E2%80%99%20%28Article%204%20%281%29%20GDPR%29.

protection impact assessments; being the focal point for the authority on any matters relating to GDPR, personal data and any other appropriate matters (Art. 38-39 GDPR).[101] A DPO may be a staff member of the controller or processor or be externally employed on the basis of a service contract. As clarified in the project proposal, DPOs of project partners will contribute and help with GDPR compliance.

In any case, it is required to have a DPO when the core activities of the controller or processor consist of processing on a large-scale of a special category of data under Art. 9 is concerned, such as health-related and biometric data.[102] The EDPB guidelines on DPOs[103] clarify core activities as "*primary activities and do not relate to the processing of personal data as ancillary activities*". Reference is made to the example of a hospital, whose primary activity is to provide healthcare, and which cannot do so without the processing of health data. Therefore, the processing of these data is a main activity of the hospital, and a DPO must be designated. By analogy, it can thus be argued that when health-related and biometric data are collected, stored and analysed for the use cases validation, the processing of these data constitutes a main activity of the controller. Second, large-scale means that the purpose of the processing activity is to process a significant amount of personal data at regional, national, or supranational levels. It may involve a large number of data subjects and is likely to present a high risk.[104] To determine what large-scale is, one can look at the number of data subjects, the volume of data, the duration or permanence of the processing activity, and the geographical extent of the processing activity. Processing activities by individual actors, on the other hand, are unlikely to be considered large-scale. Depending on the latter conditions, the data controllers in the healthcare and identity verification use cases will have to appoint a DPO, if not yet appointed. In any case, a data controller can always voluntarily appoint a DPO.

Art. 35 GDPR requires a **DPIA** "*when a type of processing in particular using new technologies, and taking into account the nature, scope, context and purposes of the processing, is likely to result in a high risk to the rights and freedoms of natural persons*". A DPIA is an assessment of the impact of the envisaged processing operations on the protection of personal data, and is part of the protection by design principle. A single assessment may address a set of similar processing operations that present similar high risks. It includes a systematic description of the envisaged processing activities and purposes, an assessment of the necessity and proportionality of the processing activities towards the purposes, an assessment of the risks to the data subjects' rights, and the measures to address the risks.[105] Art. 35(3) GDPR further elaborates on the conditions under which the assessment is required.

---

[101] Lambert, P., 2016, *The Data Protection Officer: Profession, Rules, and Role*, CRC Press, pp. 67-73.
[102] Art. 37(1)(c) GDPR requires in fact a DPO if the data controller's core activity consists of large-scale processing of special categories of data under Art. 9 or 10 GDPR.
[103] Article 29 Working Party, 2016, *Guidelines on Data Protection Officers ('DPOs') (WP 243)*, https://ec.europa.eu/newsroom/article29/items/612048.
[104] Recital 91 GDPR.
[105] Bieker, F., Martin, N., Friedewald, M. & Hansen, M., 2018, 'Data Protection Impact Assessment: A Hands-On Tour of the GDPR's Most Practical Tool', *Privacy and Identity Management*, Springer, pp. 207-220; Lievens E. & van

It may be appropriate, in the context of the use of generative AI to seek the views of those affected by the system, either the data subjects themselves or their representatives in the area of intended processing. In addition to the reviews to assess whether the DPIA is rightly implemented, regular monitoring and reviews of the risk assessments need to be carried out, since the functioning of the model may exacerbate identified risks or create new ones. Those risks are related to personal data protection, but are also related to other fundamental rights and freedoms.[106]

A DPIA is required in case of (a) a systematic and extensive evaluation of personal aspects relating to natural persons which is based on *automated processing*, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person; (b) processing on a *large scale of special categories of data* referred to in Article 9(1), or of personal data relating to criminal convictions and offences referred to in Article 10; or (c) a systematic monitoring of a *publicly accessible area on a large scale*.

Whether the healthcare and identity verification use cases will trigger the DPIA obligation under point (b) will depend on the specific details that have to become clearer in the coming months – especially through the use case questionnaires and subsequent workshop with the use case partners.

der Hof, S., 2018, 'The importance of privacy by design and data protection impact assessments in strengthening protection of children's personal data under the GDPR', *Communications Law*, pp. 33-43; Kloza, D., van Dijk, N. Casiraghi, S, Maymir, S.V., Roda, S., Tanas, A. & Konstantinou, I., 2020, 'Towards a method for data protection impact assessment: Making sense of GDPR requirements', https://doi.org/10.31228/osf.io%2Fes8bm.

[106] EDPS, 2024, *Generative AI and the EUDPR. First EDPS Orientations for ensuring data protection compliance when using Generative AI systems*, https://www.edps.europa.eu/system/files/2024-06/24-06-03_genai_orientations_en.pdf.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

6 DATA GOVERNANCE

Over the last years, the EU has adopted horizontal legislation aimed at regulating the sharing, access and use of data within its member states. Two prominent regulations within the European broader data strategy[107] are the Data Act (DA) and the Data Governance Act (DGA). Whereas the DA clarifies who can create value from data and under which conditions, the DGA regulates processes and structures that facilitate voluntary data sharing. Together, they aim to facilitate reliable and secure access to data. Given that data is a central element in various ACHILLES activities, such as the development of AI systems – both the DA and DGA are important to the project as they provide a regulatory framework to help manage data access and sharing in a compliant manner.

6.1 Data Act

The DA aims to ensure the **fair distribution of data value** by establishing rules to facilitate data access and data use within the EU. Its primary objective is to enhance data availability, promote fair access, and protect user rights while ensuring personal data protection.[108] It is a cross-sectoral legislation and applies across all sectors.

The DA lays down rules for making product data and related service data available to users[109], making data available by data holders[110] to data recipients[111], making data available by data holders to public sector bodies for the public interest, facilitating switching between data processing services, establishing safeguards against unlawful third-party access to non-personal data, and developing interoperability standards for data to be access, transferred and used (Art. 1(A)). The regulation covers both personal and non-personal data (Art. 1(2)).[112]

[107] European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and The Committee of the Regions a European strategy for data, 19 February 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066.

[108] Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act), OJ L, 2023/2854, 22.12.2023.

[109] A user is defined as "natural or legal person that owns a connected product or to whom temporary rights to use that connected product have been contractually transferred, or that receives related services" (Art. 2(12) DA).

[110] A data holder is defined as "*natural or legal person that has the right or obligation, in accordance with this Regulation, applicable Union law or national legislation adopted in accordance with Union law, to use and make available data, including, where contractually agreed, product data or related service data which it has retrieved or generated during the provision of a related service*" (Art. 2(13) DA).

[111] Data recipient is defined as "*a natural or legal person, acting for purposes which are related to that person's trade, business, craft or profession, other than the user of a connected product or related service, to whom the data holder makes data available, including a third party following a request by the user to the data holder or in accordance with a legal obligation under Union law or national legislation adopted in accordance with Union law*" (Art. 2(14) DA).

[112] European Commission, 2024, *Data Act*, https://digital-strategy.ec.europa.eu/en/policies/data-act.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

The DA concerns **data generated by the use of a product or related service**. Any data not generated by a product or related service falls outside the scope of this Act, seemingly indicating – though subject to further research and analysis of use cases – that it would not necessarily apply to all project activities. 'Connected product' is defined as "*an item that obtains, generates or collects data concerning its use or environment and that is able to communicate product data via an electronic communications service, physical connection or on-device access, and whose primary function is not the storing, processing or transmission of data on behalf of any party other than the user*" (Art. 2(5)) and 'related services' is defined as "*a digital service, other than an electronic communications service, including software, which is connected with the product at the time of the purchase, rent or lease in such a way that its absence would prevent the connected product from performing one or more of its functions, or which is subsequently connected to the product by the manufacturer or a third party to add to, update or adapt the functions of the connected product*" (Art. 2(6)).

The DA's provisions must be met throughout the ACHILLES project, particularly in the development and deployment of the ACHILLES IDE. This is crucial as Art. 1(4) specifies that when the regulation refers to 'connected products or related services', they also include **virtual assistants** insofar as they interact with a connected product or related service, which is what the ACHILLES IDE is supposed to do when giving advice on the development of AI systems. In addition, these provisions may be relevant for the AI systems being developed through the IDE. Indeed, one of the project's KPIs focuses on ensuring techniques for privacy-preserving data sharing that are compliant with the DA.

The DA lays down different obligations and requirements. Chapter II includes provisions on the ***rights of users to use data connected products and related services***.[113] Art. 3 includes the obligation to design and manufacture connected products and related service data so that they, including the metadata necessary to interpret and use those data, are easily, securely, free of charge in a comprehensive, structured, commonly used and machine-readable format, and, where relevant and technically feasible, directly accessible to the user. Art. 4 establishes the right of users to request data holders to make data available when data cannot be directly accessed by the users from the connected product or related service – based on a simple request. When adversely affecting the health, safety or security of natural persons, users and data holders may contractually restrict or even prohibit accessing, using or further sharing data. In the context of ACHILLES, when data users request access and use data generated by the user of a product or related services, the data holders are required to provide the data free of charge without delay. However, paragraph 6 specifies that trade secrets can only be disclosed if all measures have been taken to preserve their confidentiality. Although users have a right to get access to data, paragraph 10 imposes restriction on the use of this data to for developing competing products or sharing it with third parties for similar purposes. In addition, paragraph 14 allows

[113] Art. 7 DA exempts some microenterprises, small enterprises and SMEs from the obligation scope of Chapter II.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

data holders to use non-personal data generated by the use of a product or related service if they have a contractual agreement with the user.

Art. 5 further introduces the ***right of users to share data with third parties***. Upon request by users, data holders must make available readily available data to a third party. Gatekeepers, defined under the Digital Markets Act (*infra* 7.2), are not eligible as third parties. Similar rules regarding trade secrets, as contained in Art. 4 apply for this right. Art. 6 contains certain limitations to what third parties can do with the data they receive pursuant Art. 5. In the first place, third parties can only process the data made available to it for the purposes and conditions agreed with the user and in accordance with EU and national law, such as the GDPR. The third-party must erase data when no longer necessary for agreed purpose. Third parties can also not impede users' rights under Art. 5, use data to profile natural persons, make data it receives available to another third party (unless under certain conditions) or to a gatekeeper, use the data to develop a competing product with the connected product, use data so that it adversely impact security of connected product or related service, disregard the measures agreed with a data holders or with trade secret holder, and prevent the user from making data it receives available to other parties.

Chapter III includes horizontal obligations for data holders to make data available in ***business-to-business relations*** (Art. 12). First of all, Art. 8 contains the conditions under which data holders must make data available to data recipients when they are obliged to do so according to Art. 5. Data holders and data recipients must agree on arrangements, but they should in any case be under *fair, reasonable and non-discriminatory terms* and conditions and in a transparent manner. Data holders are prohibited to discriminate between comparable categories of data recipients, as well as on an exclusive basis. Unless the law provides otherwise, an obligation to make data available to a data recipient shall not oblige the disclosure of trade secrets. Article 9 specifies the way of *compensation* for making data available, i.e. non-discriminatory and reasonable. The notion of 'reasonable' compensation is open to interpretation, but the DA states that compensation may include a margin. However, the provision states the elements to take into account when determining the compensation, such as costs incurred to make data available, and investments in collection and production of data. If data recipients are an SME or not-for-profit organisation, the compensation must be limited to the costs incurred in making data available. Art. 11 foresees that data holders may apply technical protection measures, such as smart contracts and encryption, to prevent unauthorised access to data and to ensure compliance with obligations.

Chapter IV deals with the ***unfair contractual terms*** related to data access and use between enterprises. Article 13 clarifies that such terms are not binding on the disadvantaged enterprise. It clarifies when a contractual term is unfair, namely when the nature of its use grossly deviates from good commercial practice in data access and use, contrary to good faith and fair dealing (paragraphs 3-4).

Chapter V includes obligations to make ***data available to public sector bodies***, the EC, European Central Bank and Union bodies on exception needs, such as when data is needed to respond to a public emergency (Art. 14-18). Once public bodies receive data, they have to oblige certain restrictions (Art.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

19). Public bodies cannot use data incompatible with the purpose for which they were requested; they must implement technical and organisational measures to preserve the confidentiality and integrity data, erase data when no longer necessary for the purpose. Data holders are also entitled to *fair compensation* to cover technical and organisational costs and, where applicable, costs of anonymisation, pseudonymisation, aggregation and technical adaptation, and a reasonable margin (Art. 20). The rules in Chapter V seem less relevant for the ACHILLES activities.

Chapter VI foresees the rules for ***switching between data processing services***. Providers must take measures to ensure that customers of their service can switch to another data processing service covering the same service type, which is provided by a different service provider. For that purpose, data service providers shall remove commercial, technical, contractual, and organisational obstacles (Art. 23-31). Chapter VII includes rules related to ***unlawful international governmental access*** and transfer of non-personal data (Art. 32). Both the rules in Chapter VI and Chapter VII seem less relevant for the ACHILLES activities.

Lastly, Chapter VIII sets out the rules on ***interoperability of data***, data sharing mechanisms and services and common European data spaces (on the latter more, *infra* 6.3), which are important to the ACHILLES activities. A key objective of the project is to develop a framework for data standardisation and interoperability to facilitate datasets integration (WP1). Task 1.1 focuses on aligning efforts for data standardisation, in particular for the validation of the use cases in WP7, while also ensuring collaborative development and interoperability of data-sharing components within the ACHILLES IDE in WP6. These efforts will culminate in Deliverable 1 "*Guidelines for data standardisation and interoperability*". For that purpose, adherence to the relevant provisions outlined in Chapter VIII of the DA will be essential to ensure compliance in the validation of the use cases and the development of the IDE.

Art. 33 sets out the **essential requirements regarding interoperability** of data and data sharing mechanisms and services. These concern:

a) the dataset content use restrictions, licences, data collection methodology, data quality and uncertainty shall be sufficiently described to allow the recipient to find, access and use the data,
b) the data structures, data formats, vocabularies, classification schemes, taxonomies and code lists must be described in a publicly available and consistent manner,
c) the technical means to access the data, such as application programming interfaces, and their terms of use and quality of service shall be sufficiently described to enable automatic access and transmission of data between parties, including continuously, in bulk download or in real-time in a machine-readable format where that is technically feasible and does not hamper the good functioning of the connected product, and
d) the means to enable the interoperability of tools for automating the execution of data sharing agreements, such as smart contracts shall be provided.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Art. 35 includes specific provisions related to interoperability of data processing services, and Art. 36 on smart contracts for executing data sharing agreements.

6.2 Data Governance Act

The DGA seeks to facilitate the **voluntary sharing of data by individuals and businesses** and harmonise conditions for the **use of certain public sector data**. It applies cross-sectoral.[114] This is particularly relevant for the development of new products and services, including the training of AI systems.

The DGA wants to boost the development of **trustworthy data-sharing systems** through four sets of measures: (1) mechanisms to facilitate re-use[115] of public sector data that cannot be made available as open data (e.g. the re-use of health data), (2) measures to ensure that data intermediaries will function as trustworthy organisers of data sharing or pooling in Common European Data Spaces, (3) measures to make it easier for citizens and businesses to make their data available for benefit of society, and (4) measures to facilitate data sharing.[116] However, it is important to note that the DGA does not establish a new legal basis to re-use personal data under the GDPR.[117] Within the ACHILLES project, the provisions under the first and fourth measures appear to be the most relevant.

Chapter II entails the provisions on the ***re-use of certain categories of protected data by public sector bodies***. The chapter starts by explaining that it only applies to data held by public sector bodies, which are protected on the grounds of (a) commercial confidentiality, including business, professional and company secrets; (b) statistical confidentiality; (c) the protection of intellectual property rights of third parties; or (d) the protection of personal data, insofar as such data fall outside the scope of Directive (EU) 2019/1024.[118] The chapter does not apply to (a) data held by public undertakings; (b) data held by public service broadcasters and their subsidiaries, and by other bodies or their subsidiaries for the fulfilment of a public service broadcasting remit; (c) data held by cultural establishments and educational establishments; (d) data held by public sector bodies which are protected for reasons of public security, defence or national security; (e) data the supply of which is an activity falling outside the scope of the public task of the public sector bodies.

---

[114] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), OJ L 152, 3.6.2022.

[115] Re-use is defined as "*the use by natural or legal persons of data held by public sector bodies, for commercial or non-commercial purposes other than the initial purpose within the public task for which the data were produced, except for the exchange of data between public sector bodies purely in pursuit of their public task*" (Art. 2(2)).

[116] European Commission, 2024, *European Data Governance Act*, https://digital-strategy.ec.europa.eu/en/policies/data-governance-act.

[117] European Commission, 2024, *New practical guide to the Data Governance Act*, https://digital-strategy.ec.europa.eu/en/library/new-practical-guide-data-governance-act.

[118] Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast), OJ L 172, 26.6.2019.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Under ACHILLES and in the context of exchange of data held by public sector bodies, such entities could be subject to the obligations set out in this chapter.

The chapter starts by stating that ***exclusive arrangements*** on the re-use of such data are prohibited (unless under certain conditions) (Art. 4). Art. 5 sets out the **conditions for re-use**. Public sector bodies must first make publicly available the conditions for re-use and procedures to request this. These conditions must be *non-discriminatory, transparent, proportionate and objectively justified* in light of categories of data, purposes of re-use and nature of the data for which re-use is allowed. The nature of data must always be protected, e.g., by using anonymisation techniques or by being within a secure processing environment. Re-users are prohibited to re-identify any data subject to whom the data relates and must take technical and operational measures to prevent re-identification and to notify any data breach resulting in the re-identification of the data subjects concerned to the public sector body. Re-use must also be in compliance with intellectual property rights. Confidential data cannot be disclosed as a result of allowing re-use. Importantly, if public sector bodies cannot grant access to certain data to re-use, they should help seek consent to re-use the data. Art. 6 foresees the fees public sector bodies can charge to re-use of data: they must be non-discriminatory, proportionate and objectively justified and shall not restrict competition, and must be limited to the necessary costs.

Chapter II includes the requirements applicable to ***data intermediation services***. Data intermediation services are "*services which aim to establish commercial relationships for the purposes of data sharing between an undetermined number of data subjects and data holders on the one hand and data users on the other, through technical, legal or other means, including for the purpose of exercising the rights of data subjects in relation to personal data*" (Art. 2(11)).[119] Examples are data marketplaces. These services are subject to rules to ensure they function as trustworthy organisers of data sharing. They must act as neutral third parties to connect individuals and companies with data users, and facilitate data sharing between parties. However, they cannot directly use data that they intermediate for financial profit.[120] For instance, according to Art. 10, data intermediation services providers cannot use data they provide for other purposes than to put them at the disposal of data users; they must facilitate data exchange and ensure access procedures are fair, transparent and non-discriminatory; they must take measures to ensure interoperability with other data intermediation services, or; they must put in place adequate technical, legal and organisational measures in order to prevent unlawful transfer of or

[119] Excluded are: (a) services that obtain data from data holders and aggregate, enrich or transform the data for the purpose of adding substantial value to it and license the use of the resulting data to data users, without establishing a commercial relationship between data holders and data users; (b) services that focus on the intermediation of copyright-protected content; (c) services that are exclusively used by one data holder in order to enable the use of the data held by that data holder, or that are used by multiple legal persons in a closed group, including supplier or customer relationships or collaborations established by contract, in particular those that have as a main objective to ensure the functionalities of objects and devices connected to the Internet of Things; (d) data sharing services offered by public sector bodies that do not aim to establish commercial relationships.

[120] European Commission, 2024, *Data Governance Act explained*, https://digital-strategy.ec.europa.eu/en/policies/data-governance-act-explained.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

access to non-personal data and so on. And they must also notify authority before providing services (Art. 11).

Chapter IV includes the provisions related to ***data altruism***, which refers to individuals and companies giving consent or permission to make available data that they generate (voluntarily and without reward) to be used for objectives of general interest.[121] Such entities must first register in a public registry (Art. 17-19) or keep records of activities (Art. 20).

6.3 Common European Data Spaces

As part of its broader European strategy for data, the EU is trying to establish Common European Data Spaces (CEDS). These spaces aim to create an **internal market for data**, ensuring more data is made available for access and re-use in a **trustworthy and secure environment** for the benefit of European businesses and citizens.[122] The EU is currently in the process of rolling out the Common European Data Spaces across fourteen sectors: agriculture, cultural heritage, energy, finance, green deal, health, language, manufacturing, media, mobility, public administration, research and innovation, skills and tourism. The sector-specific data spaces are to be established tailored to the needs and characteristics of the sector.[123] The data spaces will complement existing legislation, such as the interoperability requirements in Art. 33 DA.

Common data spaces are meant as ***open digital environments with secure, privacy-preserving infrastructures to pool, access, share, process and use data***. They should have fair, transparent, proportionate and non-discriminatory access rules while still respecting EU legislation. They will allow data holders to grant access and share certain (non-)personal data in a controlled and compliant manner.[124]

In March 2025, the European Health Data Space (EHDS) became the first domain-specific common European data space.[125] The EHDS establishes a common framework for the use and exchange of electronic health data across the EU. It enables individuals to access, control and share their data

[121] European Commission, 2024, *Data Governance Act explained*, https://digital-strategy.ec.europa.eu/en/policies/data-governance-act-explained.
[122] European Commission, 2025, *Common European Data Spaces*, https://digital-strategy.ec.europa.eu/en/policies/data-spaces.
[123] European Commission, 2022, *Staff working document on data spaces*, https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces.
[124] European Commission, 2025, *Common European Data Spaces*, https://digital-strategy.ec.europa.eu/en/policies/data-spaces.
[125] Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847, OJ L, 2025/327, 5.3.2025.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

across borders for healthcare delivery, and enables secure and trustworthy re-use of health data for research, innovation, policy-making and regulatory activities.[126]

It will be essential to monitor the development of the Common European Data Spaces for the ACHILLES project. The EHDS, for instance, will be relevant for the healthcare use case, as well as future AI systems developed through the IDE for this domain. Beyond healthcare, the broader implementation of sector-specific data spaces must be closely observed to assess how they may facilitate data access and use within ACHILLES.

[126] European Commission, 2025, *European Health Data Space Regulation (EHDS)*, https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

7 INFORMATION SOCIETY SERVICES

In addition to personal data protection and data governance frameworks, EU horizontal legislation has been introduced regulating information society services, which may be relevant to the ACHILLES project. Given that these services **facilitate access to data**, a critical factor for designing, validating and training AI models and systems, this chapter examines two key regulations: the Digital Services Act (DSA) and the Digital Markets Act (DMA).

Both the DSA and DMA are part of the EU's digital services package and contribute to the broader European digital strategy.[127] Their shared objective is to create a secure and transparent digital environment that protects users' fundamental rights and ensures a level playing field for businesses, but also promotes innovation, growth and competitiveness.[128] Whereas the DSA focuses on regulating online intermediaries and platforms, the DMA focuses on regulating online gatekeepers.

7.1 Digital Services Act

The DSA regulates **online intermediaries and platforms**, such as marketplaces, social networks, content-sharing platforms, app stores and online travel and accommodation platforms.[129] The DSA aims to prevent illegal and harmful activities online, and the spread of disinformation, as well as ensure user safety, fundamental rights and a fair and open online platform environment.[130] It builds on the idea that everything illegal offline should be illegal online, too.

In the context of ACHILLES, the intermediation services provided by the IDE may need to be evaluated in light of the obligations set out in the DSA. Depending on its functionalities (which must become clearer along the project timeline), the IDE could qualify as an online platform, making it subject to DSA requirements concerning transparency, content moderation, and recommender systems. In addition, AI systems developed through the IDE may also need to comply with relevant DSA provisions.

The DSA includes rules for intermediary companies that connect users with content, products and services in the EU single market – whether or not they are established in the EU. This includes online marketplaces, social networks, content-sharing platforms, and online travel and accommodation

[127] European Council, 2022, *Digital services package*, https://www.consilium.europa.eu/en/policies/digital-services-package/; European Commission, 2025, *The Digital Services Act package*, https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package.
[128] European Commission, 2022, *European Commission digital strategy Next generation digital Commission*, https://commission.europa.eu/document/download/70703206-2592-4175-b10d-12f97382094a_en?filename=C_2022_4388_1_EN_ACT.
[129] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), OJ L 277, 27.10.2022.
[130] European Commission, 2022, *The Digital Services Act*, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

platforms.[131] The obligations vary depending on the type of online service providers, their role, size and overall impact on the online ecosystem, as illustrated in Figure 4. The following sections will outline the types of online service providers and their corresponding obligations, after which an analysis of how these provisions can apply to the ACHILLES IDE will follow.

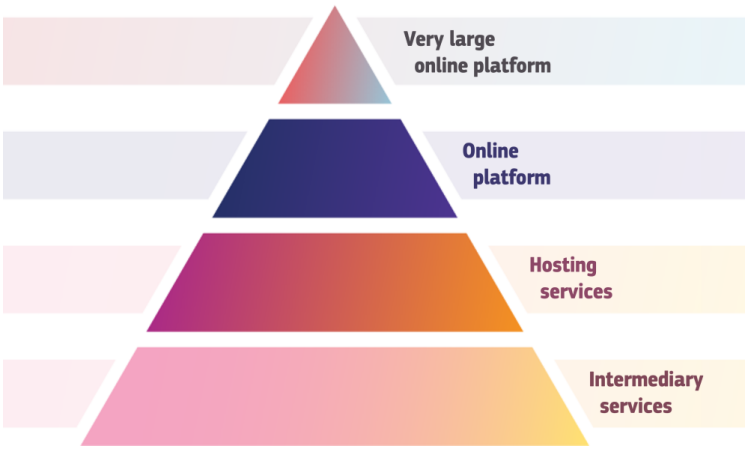

Figure 4 – DSA layers of application (European Commission 2024)132

The first set of rules applies to **all intermediary services**. "Intermediary services" are defined as any of the following information society services:[133]

i. a 'mere conduit' service, consisting of the transmission in a communication network of information provided by a recipient of the service, or the provision of access to a communication network;
ii. (ii) a 'caching' service, consisting of the transmission in a communication network of information provided by a recipient of the service, involving the automatic, intermediate and temporary storage of that information, performed for the sole purpose of making more efficient the information's onward transmission to other recipients upon their request;
iii. (iii) a 'hosting' service, consisting of the storage of information provided by, and at the request of, a recipient of the service.

For providers of intermediary services, Section 1 of the DSA outlines obligations to ensure accountability and transparency. These obligations include the requirement to designate a single point of contact to communicate with authorities (Art. 11), a single point of contact for recipients of the

[131] European Council, 2025, *Digital Services Act*, https://www.consilium.europa.eu/en/policies/digital-services-act/#act.
132 European Commission, 2024, *The Digital Services Act*, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en.
[133] 'Information society service' is defined as a service "*any service normally provided for remuneration, at a distance, by electronic means and at the individual request of a recipient of services*" (Art. 3(a) DSA j Art. 1(1)(b) Directive (EU) 2015/1535).

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

service (Art. 12), and a legal representative in case they are not established in the EU (Art. 13). Service providers must include clear information in their terms and conditions on any restrictions to the use of their services to recipients (Art. 14). They are also subject to transparency reporting obligations in the form of making publicly available a clear and comprehensible report of any content moderation that they did (Art. 15).

A second layer of rules, outlined in Section 2, applies to all **hosting services**, which is a service to store information provided by a recipient of the service (Art. 3(g)(iii)). Recital 13 clarifies that online platforms, such as social networks or online platforms allowing consumers to conclude distance contracts with traders, are considered hosting services when they store or disseminate information at the request of users. However, platforms are not classified as hosting services if the dissemination is a minor, ancillary feature of a different service – such as a comment section on a news site, which supports the main function of publishing news. In contrast, platforms like social networks, where storing and sharing comments is a core service, are considered hosting services. In addition, cloud computing and web-hosting services, which mainly provide infrastructure for applications, are excluded from the scope of hosting services if dissemination to the public is a minor feature.

Providers of hosting services must put in place mechanisms to allow individuals to *notify them of illegal content* (Art. 16). They must also provide *information* to individuals when there are any restrictions imposed on the information provided by the recipients when it is illegal or incompatible with their terms and conditions (Art. 17). They also inform the law enforcement authorities when they become aware of information giving rise to suspicion that a criminal offence involving a threat to the life or safety to an individual has taken place (Art. 18).

A third layer of rules applies to **online platforms**, defined as "*a hosting service that, at the request of a recipient of the service, stores and disseminates information to the public, unless that activity is a minor and purely ancillary feature of another service or a minor functionality of the principal service and, for objective and technical reasons, cannot be used without that other service, and the integration of the feature or functionality into the other service is not a means to circumvent the applicability of this Regulation*" (Art. 3(i)). Providers of online platforms must provide recipients of the service an effective complaint-handling system to *lodge complaints against* decisions taken by them, such as decisions on removing or disabling access or visibility of information, termination of services or recipient's account, about monetisation information (Art. 20). Recipients have the right to opt for out-of-court dispute settlement bodies to resolve disputes relating to those decisions (Art. 21). Providers must prioritise notices from trusted flaggers (Art. 22). Providers have the right to suspend services to recipients who frequently provide manifestly illegal content (Art. 23). Providers must abstain from designing online interfaces that deceives or manipulates recipients of services (Art. 25), and they must ensure that advertisements on their online interfaces are clearly presented as advertisements (Art. 26). In case providers make use of recommender systems, they must set out in their terms and conditions, the main parameters used in this system and any options for recipients to modify or influence these

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

(Art. 27). In addition, providers must implement measures to protect children's privacy, safety and security (Art. 28).

Lastly, the fourth layer of rules applies to **very large online platforms** (VLOPs) and **very large online search engines** (VLOSEs),[134] which are deemed to pose high societal risks due to the dissemination of illegal and harmful content, including disinformation.[135] VLOPs and VLOSEs are subject to carrying out a risk assessment: they must identify, analyse and assess any systemic risks in the EU following the design or functioning of their services and system or use made of their services (Art. 32) and must take appropriate mitigation measures (Art. 33). They must also allow independent audits to assess compliance with obligations (Art. 37). The audits serve to evaluate whether and to what extent online platforms and search engines comply with obligations and commitments made in codes of conduct and crisis protocols. To ensure efficient audits, these platforms and engines should provide access to all relevant data (Art. 40). In that sense, Art. 40 can be seen as introducing a new right to access data. The obligation includes providing information on their algorithmic systems' design, logic, operation and testing. In addition, vetted researchers should be granted access to data when they want to conduct research that contributes to the detection, identification and understanding of systemic risks in the EU, and to the assessment of the adequacy, efficiency and impacts of the risk mitigation measures (Art. 40(4)).

In the context of ACHILLES, it will be necessary to assess whether the ACHILLES IDE qualifies as an online platform under the DSA. This classification depends on the specific features and functionalities of the IDE. The envisaged IDE will be designed to host and process AI models, datasets and compliance tools, but most likely not primarily intended to store and disseminate content to the public, which is a key characteristic of online platforms under the DSA. If the IDE is intended for private or restricted use, such as for developers within an organisation, it is less likely to fall within the DSA's definition of an online platform. However, if the IDE allows for the public sharing or dissemination of AI models, datasets or compliance advice or enables interaction with the public, it may be classified as an online platform. In that case, the IDE is subject to obligations on content moderation, transparency, and ensuring accountability in terms of how data and content are managed. For instance, the IDE may need to provide a complaint-handling system for users and ensure transparency regarding how AI models and datasets are shared or disseminated.

[134] Art. 33 clarifies VLOPs and FLOSEs are online platforms or online search engines which have a number of average monthly active recipients of the service in the Union equal to or higher than 45 million, and which are designated as very large online platforms or very large online search engines by the European Commission.
[135] In meantime, the European Commission has designated the VLOPs and VLOSEs, including AliExpress, Amazon Store, Apple Store, Pornhub, Booking.com, Google Search, Google Play, Google Maps, YouTube, Shein, LinkedIn, Facebook, Instagram, Bing, TikTok, X, Wikipedia, Zalando.
European Commission, 2025, *Supervision of the designated very large online platforms and search engines under DSA*, https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Furthermore, as future AI systems are developed through the IDE, each system may require a separate assessment to determine whether it triggers the DSA's scope. If the systems are used as platforms to share or disseminate content to the public, they may be subject to the DSA.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

7.2 Digital Markets Act

The DMA aims to create markets in the digital sector that are fairer and more contestable.[136] It applies to core platform services provided or offered by **gatekeepers** to business users established in the EU or end users established or located in the EU (Art. 1(2)). The DMA regulates gatekeepers, referring to undertakings providing core platform services, such as online intermediation services; online search engines; online social networking services; video-sharing platform services; number-independent interpersonal communications services; operating systems; web browsers; virtual assistants; cloud computing services; or online advertising services (Art. 2(2)), designated by the European Commission (Art. 3). It concerns undertakings that have a *significant impact on the internal market*, provides a core platform services important gateway for business users to reach end users, and enjoys and entrenched and durable position in its operations or is likely to enjoy such position in future. Over time, the European Commission has designated some gatekeepers, including Alphabet, Amazon, Apple, Booking, ByteDance, Meta and Microsoft.[137]

Gatekeepers are subject to a range of obligations and prohibitions (Art. 5-15). They must allow third parties to *inter-operate* with the gatekeeper's own services in certain specific situations; allow their business users to *access the data that they generate* in their use of the gatekeeper's platform; provide companies *advertising* on their platform with the tools and information necessary for advertisers and publishers to carry out their own independent verification of their advertisements hosted by the gatekeeper); allow their business users to *promote* their offer and conclude contracts with their customers outside the gatekeeper's platform. In addition, they are prohibited from treating services and products offered by the gatekeeper itself more favourably; preventing consumers from linking up to businesses outside their platforms; preventing users from un-installing any pre-installed software or app if they wish so; tracking end users outside of the gatekeepers' core platform service for the purpose of targeted advertising, without effective consent having been granted.

The DMA also introduces a ***right to access to data*** to address the issue of unfair data practice. Gatekeepers must provide business users operating on the core platform services of gatekeepers a right to access data. Gatekeepers must provide access to aggregated and non-aggregated data, including personal data, related to products and services offered by those business users on core platforms services (art. 6(10)). Upon request by business users, free, efficient, qualitative and continuous real-time access must be provided. Access is limited to personal data directly related to customers of business users. In addition, gatekeepers must grant third-party providers of online search engines, upon their request, access to the anonymised data on rankings, searches, clicks and views

[136] Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act),
OJ L 265, 12.10.2022.
[137] European Commission, 2025, *Gatekeepers*, https://digital-markets-act.ec.europa.eu/gatekeepers_en.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

related to end users' free and paid searches on those online search engines (Art. 6(11)). Access will be on fair, reasonable and non-discriminatory terms, the so-called FRAND terms.[138]

In the context of ACHILLES, it seems unlikely that the IDE will be considered as a gatekeeper under the DMA due to the strict conditions. The IDE is not intended to be a core platform service providing essential gateways for business users to access end users at the scale typically required to meet the DMA's threshold for gatekeeper status. Therefore, the DMA obligations are unlikely to apply to the IDE. However, this assessment may evolve depending on future developments in the project, particularly if the IDE's features or scope change in the sense that it could influence its status.

[138] FRAND is an acronym for "fair, reasonable and non-discriminatory".

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

8 CYBERSECURITY

In addition to the requirements set out in Art. 15 of the AI Act, which imposes obligations regarding cybersecurity on AI systems (*supra* 4.1.3.6), there are other specific legislative frameworks on cybersecurity that are relevant to the ACHILLES project, namely the NIS2 Directive, the EU Cybersecurity Act and the EU Cyber Resilience Act. These frameworks offer more detailed and concrete provisions that complement the cybersecurity requirements in the AI Act. The frameworks are particularly important since ensuring robust cybersecurity is critical for the secure development of AI models and systems, given the increasing number of cyber threats in the context of automation.

8.1 NIS2 Directive

The EU Directive on measures for a high common level of cybersecurity across the Union, known as the NIS2 Directive,[139] replaced the NIS1 Directive and entered into force in January 2023. As a minimum harmonisation directive, Member States can impose stricter cybersecurity obligations in their jurisdictions (Art. 5).

NIS2 imposes cybersecurity requirements on entities that operate in critical and important sectors. Besides imposing obligations that require Member States to adopt national cybersecurity strategies and supervisory and enforcement obligations on Member States, NIS2 also lays down **cybersecurity risk-management measures and reporting obligations** for certain entities, as well as rules and obligations on **cybersecurity information sharing** (Art. 1).

The Directive applies to entities that carry out their activities or provide services in an EU member state, if they are considered mid-sized or large organisations, and if they operate in one of the **critical sectors** included in the Annexes. NIS2 applies to certain types of entities specified in Annex I (entities active in sectors of energy, transport, banking, financial market infrastructure, health, drinking water, waste water, digital infrastructure, ICT service management, public administration, and space) and Annex II (entities active in sectors of postal and courier services, waste management, chemicals, food, manufacturing, digital providers, research). Annex I categorises entities as either '*Essential*' or '*Important*' based on their annual revenue and size. Both types of entities must comply with the same security measures, but Essential entities are subject to continuous proactive supervision, while Important entities are monitored reactively, typically after an incident or case of non-compliance.[140]

[139] Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive), OJ L 333, 27.12.2022.
[140] KPMG, 2023, *Network & Information Security Directive (NIS2) NIS2 (EU) Directive Readiness - Levelling-up your IT and OT security capabilities in light of the NIS2*, https://assets.kpmg.com/content/dam/kpmg/pl/pdf/2023/10/kpmg-network-and-information-security-directive-nis2.pdf; NIS2 Compliant.org, 2024, *Comprehensive guide to the NIS 2 directive V.2.0*, https://nis2compliant.org/wp-content/uploads/2024/07/NIS-2-guide-1.pdf.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Under Chapter IV, NIS2 imposes several **obligations** on entities. It requires entities to implement a comprehensive *proactive risk management*, ensuring the adopting of appropriate industry-standards cybersecurity measures (Art. 21). Entities are also required to *report any incident* that has a significant impact on the provisions of their services: they should issue without undue delay (within 24h) an early warning, an incident notification (within 72h), issue upon request an *intermediate report* on the status of incident and crisis management, and submit a final report on incited, mitigation measures and effects (Art. 23).

To determine whether NIS2 is relevant for ACHILLES, it must be assessed whether the entities involved provide a critical service or essential function directly to an end client or key supplier that could impact public safety or economic stability, and whether the entities operate in a sector covered by Annex I or Annex II. Regarding the ACHILLES IDE itself, it may trigger NIS2 applicability if considered a digital provider, as referred to in point 6 of Annex II. However, upon further inspection, this sector concerns providers of online marketplaces, providers of online search engines and providers of social networking services platforms. As discussed (*supra* 7), the IDE is unlikely to be considered an online search engine, an online marketplace or a social networking services platform – although this should be reassessed based on further project developments. As for the healthcare use case, Annex I, point 5 specifies that entities active in the sector of health fall under NIS2 when being a mid-sized or large organisation. However, when taking a closer look, the scope is delineated in the sense that only certain types in the sector fall under its scope. It concerns healthcare providers, EU reference laboratories, or entities carrying out R&D activities of medicinal products or manufacturing pharmaceutical products or medical devices critical for public health emergencies. This does not appear to apply directly to the use case but should be further explored in the coming months of the project.

8.2 EU Cybersecurity Act

Another cybersecurity-related legislative framework is the EU regulation on information and communications technology cybersecurity certification, known as the EU Cybersecurity Act.[141] It establishes a new and **stronger mandate** for the EU agency for cybersecurity, i.e. ENISA, and introduces a framework for **voluntary European cybersecurity certification schemes** for ICT products, services and processes.[142] The latter is to build trust, increase growth in the cybersecurity market and facilitate trade across the EU.

While the EU Cybersecurity Act provides a framework for voluntary certification, it does not impose specific cybersecurity requirements that are directly relevant to ACHILLES.

[141] Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act), OJ L 151, 7.6.2019.
[142] European Council, 2024, *How the EU is strengthening its cybersecurity*, https://www.consilium.europa.eu/en/policies/cybersecurity/#act.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

8.3 EU Cyber Resilience Act

A final cybersecurity-related framework to consider is the EU Regulation on horizontal cybersecurity requirements for products with digital elements and amending Regulations, known as the EU Cyber Resilience Act.[143] It entered into force in December 2024, but the main obligations will apply from December 2027 only.

The Act imposes **conditions for the development of secure hardware and software ICT products**. It aims to ensure that manufacturers take security measures throughout products' lifecycle, and conditions to allow users to take cybersecurity into account when selecting and using digital products.[144]

It imposes new cybersecurity requirements on manufacturers and retailers concerning the planning, design, development, and maintenance of products. The regulation applies to all products that are connected, either directly or indirectly, to another device or network except for specified exclusions such as certain open-source software or services products that are already covered by existing rules, which is the case for medical devices, aviation and cars.[145] In particular, Art. 2(2)(a) explicitly ***excludes*** products with digital elements covered by the Medical Devices Regulation, which may be relevant for the healthcare use case, but also for future AI systems in this domain being developed through the ACHILLES IDE.

A product with digital elements is defined as "*software or hardware product and its remote data processing solutions, including software or hardware components being placed on the market separately.*" In turn, 'remote data processing' is defined as "*data processing at a distance for which the software is designed and developed by the manufacturer, or under the responsibility of the manufacturer, and the absence of which would prevent the product with digital elements from performing one of its functions*" (Art. 3(1)(2)).

These products can only be made available on the market when they meet a set of **essential cybersecurity requirements**, such as having a secure default configuration, without known exploitable vulnerabilities, mitigation measures for vulnerabilities, control mechanisms for unauthorised access, mechanisms to ensure confidentiality, integrity and availability and so on (Art. 6 j Part I of Annex I). The processes put in place by the manufacturer must also comply with the vulnerability handling requirements set out in Part II of Annex I that focus on identifying and

[143] Regulation (EU) 2024/2847 of the European Parliament and of the Council of 23 October 2024 on horizontal cybersecurity requirements for products with digital elements and amending Regulations (EU) No 168/2013 and (EU) 2019/1020 and Directive (EU) 2020/1828 (Cyber Resilience Act), OJ L, 2024/2847, 20.11.2024.
[144] European Commission, 2022, *Cyber Resilience Act*, https://digital-strategy.ec.europa.eu/en/library/cyber-resilience-act.
[145] European Commission, 2022, *Cyber Resilience Act*, https://digital-strategy.ec.europa.eu/en/library/cyber-resilience-act.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

documenting vulnerabilities and risks, applying regular tests and reviews of security, taking measures to facilitate sharing information about vulnerabilities, and ensuring security updates.

Some products are considered important products and made subject to conformity assessment procedures (Art. 7(1) j Annex III j Art. 32), including password managers, software searching for malicious software, VPN, network management systems, boot managers, operating systems, smart home general purpose virtual assistants and so on. Also included are identify management systems and privileged access management software, including authentication and access control readers, including biometric readers, digital certificate issuance software, and physical or virtual network interfaces. However, paragraph 2 details that being included in Annex III does not automatically render the product subject to the conformity assessment procedure. It requires that the product must be either (a) a product that primarily performs functions critical to the cybersecurity of other products, networks or services, including securing authentication, or (b) the product performs a function that carries significant risks of adverse effects in terms of its intensity and ability to disrupt, control or cause damage to a large number of other products, health, security, safety of its users through direct manipulation, such as central system function.

Lastly, Art. 12 of the EU Cyber Resilience Act establishes how it ***interacts*** with Art. 15 of the AI Act. Art. High-risk AI systems under Art. 6 AI Act will be deemed to comply with the cybersecurity requirements set out in Art. 15 AI Act if (1) the products fulfil essential cybersecurity requirements in Part I of Annex I, (2) the processes put in place by manufacturer comply with essential cybersecurity requirements in part II of Annex II, and (3) level of cybersecurity protection required under Art. 15 AI Act is demonstrated in the EU declaration of conformity issued under the Cyber Resilience Act.

While the IDE itself may not fall under the scope of the EU Cyber Resilience Act, as it is not primarily focused on critical cybersecurity functions, the products and AI systems developed through the IDE may be subject to compliance. For instance, any AI systems classified as high-risk under the AI Act could trigger the need to meet the cybersecurity requirements in the EU Cyber Resilience Act. These requirements include ensuring secure default configurations, managing vulnerabilities, and implementing processes for vulnerability handling, all of which are crucial for developing AI systems that are resilient to cyber threats. In addition, the EU Cyber Resilience Act mandates manufacturers to provide essential cybersecurity protections, which may impact AI systems developed through the IDE that involve remote data processing, as these systems will need to comply with the relevant security standards.

The EU Cyber Resilience Act also introduces conformity assessment procedures for certain high-risk products, which could apply to AI systems if they are deemed to perform critical functions in cybersecurity or if their failure could disrupt public safety, health, or security. As AI systems evolve and become integral to sectors like healthcare, their alignment with the EU Cyber Resilience Act's requirements will be essential for compliance, particularly for ensuring that they can be safely deployed and used within the EU.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

9 SECTORAL LEGISLATION

Besides EU horizontal law, sectoral legislation will also play a role in the ACHILLES project. In particular, the AI systems developed through the IDE will sometimes trigger sectoral law. It will, therefore, be important to explore and assess further which sectoral laws are relevant and to what extent the IDE will have to provide compliance information in relation to these laws.

At this stage, identifying potential sector-specific legislation is essential in setting the groundwork for more detailed compliance evaluations later in the process. To facilitate this, KUL will prepare a **questionnaire** to gather more information from use case owners in the coming months. This questionnaire will help in understanding the legal requirements specific to each use case. In addition, a workshop will be organised to discuss them in more depth. As AI systems are developed through the ACHILLES IDE, these specific use cases will trigger the application of sectoral law.

9.1 Medical Devices Regulation

Regulation (EU) 2017/745 on medical devices, also known as the Medical Devices Regulation (MDR), sets out rules for the placing on the market, making available on the market or putting into service of medical devices for human use and accessories for such devices in the EU.[146] The MDR could be relevant to use cases concerning healthcare and pharmaceutics. The ACHILLES healthcare use case focuses on developing, deploying and analysing automated ML-based diagnostic tools, assessing the detection and grading of ophthalmological disease conditions from eye fundus scans and complementary clinical data. It specifically concerns applications like Age-related Macular Degeneration (AMD) diagnosis, Diabetic Retinopathy (DR) screening, and Glaucoma detection. Similarly, the HERA use case in the pharmaceutical sector seeks to enhance efficiency and regulatory compliance by developing AI-driven systems such as a context-aware search engine, automated compliance monitoring, document summarisation and personalised AI assistants for pharmaceutical workers.

The MDR enhances controls to ensure the **safety and effectiveness of medical devices**.[147] Article 51(1) MDR classifies medical devices into four main classes based on their intended purposes and risks: Class I (low risk), Class IIa (medium risk), Class IIb (medium to high risk), and Class III (high risk).

[146] Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC, OJ L 117, 5.5.2017.

[147] Medical devices are "*any instrument, apparatus, appliance, software, implant, reagent, material or other article intended by the manufacturer to be used, alone or in combination, for human beings for one or more of the following specific medical purposes: diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of disease, [...] injury or disability; investigation, replacement or modification of the anatomy or of a physiological or pathological process or state; providing information by means of in vitro examination of specimens derived from the human body, including organ, blood and tissue donations*" (Art. 2(1)).

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Software intended to assist in diagnostic or therapeutic decisions is classified based on its impact. Rule 11 of Annex VIII of the MDR clarifies that software designed to assist in diagnostic or therapeutic decisions may fall into Class IIa, unless it affects critical health outcomes, in which case it is classified as Class III (high risk). In cases where the software is designed to monitor vital physiological parameters that could result in immediate danger to the patient, it would be classified as Class IIb. Other software not directly involved in these critical functions may fall under Class I.

It has been argued that ***AI technologies*** can indeed be classified as high-risk medical devices under the MDR when medicinal product incorporated into the device has an *ancillary* role concerning the device. [148] Manufacturers of medical devices incorporating AI must comply with a range of requirements, such as establishing a risk management process, keeping technical documents and EU declaration of conformity, drawing up a quality management system, reporting incidents, conducting a conformity assessment and so on. The application of these requirements will depend on the specific use cases and the classification of the AI systems involved.

In addition to the MDR, it will also be essential to consider other relevant healthcare and pharmaceutical legislation that may be triggered by the use cases and AI systems developed through the ACHILLES IDE. This includes Regulation (EC) No 726/2004 on *authorisation and supervision* of medicinal products for human and veterinary use, Regulation (EU) No 536/2014 on *clinical trials* on medicinal products for human use, or Directive 2011/24/EU on the application of patie*nts' rights in cross-border* healthcare. Further examination of these regulations will be necessary to ensure full compliance with sector-specific legal frameworks.

9.2 Intellectual property rights legislation

The EU AI Act already contains some specific obligations with respect to intellectual property protection. This could be relevant to certain use cases within ACHILLES, such as the SCRIPTA and HERA use cases, as well as the ACHILLES IDE itself. SCRIPTA, designed to identify and correct issues in script generation models, must ensure that content aligns with editorial, ethical, and legal benchmarks and safeguards against the misuse of AI technologies. It aims to test literary works to ensure their moral, ethical and legal integrity. Similarly, the HERA use case involves an LLM-based virtual assistant, and can touch upon the intersection of AI and IP law. Both use cases raise questions not only about IP-protected works to train AI models but also about the potential for AI systems to generate content that could infringe on IP rights.[149]

[148] Bitkina, O., Park, J. & Kim, H., 2023, 'Application of artificial intelligence in medical technologies: A systematic review of main trends', *Digital Health;* Mennella, C., Maniscalco, U., De Pietro, G. & Esposito, M., 2024, 'Ethical and regulatory challenges of AI technologies in healthcare: A narrative review', *Heliyon*; Onitiu, D., Wachter, S. & Mittelstadt, B., 'How AI Challenges the Medical Device Regulation: Patient Safety, Benefits, and Intended Uses'.
[149] Vanherpe, J., 2024, 'Artificial Intelligence and Intellectual Property Law', *The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence*, pp. 211-227.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

The **AI Act** specifically addresses IP concerns related to GPAI models. For such models, providers must draw up and make publicly available a sufficiently detailed summary of the content used for training the GPAI model (Art. 53(1)(c) and recitals 105-109).

Beyond the AI Act, several other EU sectoral legislation might come into play when dealing with IP protection in the context of AI. Important to notice, copyright law in the EU is primarily governed at national level. Nevertheless, there are some important EU directives. Given their nature, they must be transposed to national law by the EU Member States.

In the first place, EU **copyright** directives include relevant rules.[150] One such directive is the Directive on Copyright and Related Rights in the Digital Single Market,[151] which includes provisions relevant to AI applications. It defines when the exceptions for text and data mining applies, allowing AI systems to train on IP-protected works. It also imposes new obligations on online content-sharing service providers, requiring them to obtain authorisation of rightsholder when they allow public to access works that is protected by copyright and uploaded by their users.[152] These provisions can be relevant for ACHILLES for the AI systems that make use of copyrighted materials in their development and operation.

In addition to copyright, other EU directives protect **trade secrets** and **databases**, which are also important in the context of automation and AI. The Directive 2016/943 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure[153] safeguards against the unlawful acquisition, use, and disclosure of trade secrets. This is relevant for AI systems that rely on proprietary algorithms, models or data that need to be protected frpm unauthorised access.[154] As for databases, the Directive 96/9/EC on the legal protection of databases[155] grants a sui generis right to protect databases, especially those that involve substantive

[150] European Commission, 2024, *The EU copyright legislation*, https://digital-strategy.ec.europa.eu/en/policies/copyright-legislation.
[151] Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC, OJ L 130, 17.5.2019.
[152] FOD Economie Belgium, 2022, *European Directive on copyright and related rights in the Digital Single Market – transposition in Belgian law*, https://economie.fgov.be/en/themes/intellectual-property/intellectual-property-rights/copyright-and-related-rights/copyright/european-directive-copyright.
[153] Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, OJ L 157, 15.6.2016.
[154] European IP Helpdesk, 2021, *Trade Secrets: Managing Confidential Business Information*, https://op.europa.eu/en/publication-detail/-/publication/5f1c6d8a-f015-11eb-a71c-01aa75ed71a1/language-en; EU, 2025, *Database protection*, https://europa.eu/youreurope/business/running-business/intellectual-property/database-protection/index_en.htm.
[155] Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, OJ L 77, 27.3.1996.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

investments in data collection, verification or processing.[156] This protection could be important for AI systems that rely on large, curated databases, particularly when these databases contain IP-protected content.[157]

In conclusion, the intersection of AI development and IP protection is complex as it is regulated in different frameworks, such as the AI Act, copyright law, trade secret protection and database rights. However, the frameworks are relevant to consider in ACHILLES to ensure that IP rights of creators and innovators are respected when AI systems are developed and IP-protected content is used for training and operation.

[156] Databases are defined as "*collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means*" (Art. 1(2) Directive 96/9/EC).
[157] EU, 2025, *Database protection*, https://europa.eu/youreurope/business/running-business/intellectual-property/database-protection/index_en.htm.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

10 ETHICS CONSIDERATIONS

Besides the legal frameworks and requirements triggered by the project activities in ACHILLES, ethical considerations are also to be taken into account, especially given the involvement of human participants and the development of an AI system, i.e. the IDE itself, and the development of other AI systems through the IDE. D4.4 "*Ethics Guidelines*", led by ETICAS, already covers some ethical discussions with its aim being to identify ethical guidelines for any ethical issues that must be addressed during the project's lifecycle. It focuses on responsible research, aspects of governance and protection, informed consent, and some best practices for dissemination activities. Moreover, D4.3 will specifically focus on doctrinal research on the legal and ethical implications of designing trustworthy and privacy-preserving AI systems, and the impact assessment of the legal and ethical risks associated with the project's results. So, ethical considerations will be explored in more detail through these deliverables.

However, this chapter in D4.1 still wants to explore some ethics considerations not (yet) covered in the mentioned deliverables. It does so by focusing on ethical consent for research participants, privacy and confidentiality, facial recognition issues, algorithmic biases and issues related to hallucinations and generative AI. The last part is dedicated to discussing trustworthiness, as one of the main objectives of ACHILLES is to contribute to the trustworthy AI debate. For that reason, reference is made to the High-Level Expert Group AI Principles, given its focus on trustworthiness and the fact that the AI Act is supposed to be based on these principles.

10.1 Introduction

Although there is no universal understanding of what ethics mean, for feasibility reasons, within the project, it is narrowed down to underlying *European human normative norms*, foundational to and in the background of law. Whereas law can be considered the practical advancement of rules, ethics can be considered as the foundation on which rules are based. Ethics concerns itself with what is normatively good or bad, and with the morality of human action.[158] Ethics compliance can be seen as a means and articulation of values of our society and can go beyond legal compliance. In general, ethics compliance is important to ensure the integrity of research.[159] Legal obligations might be silent on certain research practices, such as obtaining consent for participation in scientific activities.[160] In this way, it can be an added value towards the legal obligations. Although most ethical principles are not

[158] Bickenbach, J., 2012, *Ethics, law and policy*, Thousand Oaks, pp. 19-24; Floridi L. & Taddeo, M., 2016, 'What is Data Ethics?', *Phil. Trans. R. Soc. A*; Rochel, J., 2021, 'Ethics in the GDPR: A Blueprint for Applied Legal Theorye, *International Data Privacy Law,* pp. 209-223; Whittlestone, J., Nyrup, R., Alexandrova, A. & Cave, S., 2019, 'The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions.', *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 195–200.
[159] Serio, M. e.a., 2023, 'Ethics in Legal Research', *Ethics in Research*. Springer, ALLEA, 2023, *The European Code of Conduct for Research Integrity*, https://allea.org/wp-content/uploads/2023/06/European-Code-of-Conduct-Revised-Edition-2023.pdf; European Commission, 2010, 'European Textbook on Ethics in Research', pp. 11-32.
[160] European Commission, 2010, 'European Textbook on Ethics in Research', pp. 11-32.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

legally enforceable, they are often embedded in legal instruments and thus enforceable through this route. In any case, from the perspective of good practices and credibility, they should always be taken into account.

In line with the European fundamental ethical principles governing scientific research,[161] ethical standards such as privacy, confidentiality, security, and integrity will be closely monitored during the project. Other ethical values have to be reflected in the project too, ranging from informed consent and transparency to fairness, accountability and the promotion of human dignity.[162]

10.2 Ethical consent for research participants

In WP7, human participants are involved in validating the use cases. Whenever humans participate in research activities, not only consent regarding the processing of personal data is required (*supra* 5.1.2), but also **ethical consent** to ensure that research is conducted ethically. Consent follows from more fundamental moral principles, such as respect for autonomy, agency and dignity of persons.[163]

To obtain an ethical informed consent, a simple 'yes' is insufficient. Three elements must be present: **information, voluntariness and competence**.[164] First, valid consent requires adequate information. Research subjects must be informed on the purpose of the research, funding, and risks. Not everything has to be disclosed as excessive information, so-called information overload, might result in the fact that participants can no longer absorb and understand information. Specifically, information sheets should be drafted in an understandable manner, not containing too much technical language, in the language of the participant, and allowing questions to be asked. Second, the consent must be voluntary or free. This means no manipulation, coercion, undue inducements, penalties or false promises. Third, it is required that the participant has sufficient competence to consent, referring to mental competence and capacity to understand and retain relevant information about the research. In ACHILLES, no vulnerable persons are involved in the use case validation activities. In any case, information sheets can be used to provide more information, such as a clear explanation of the aim and overall purpose, the methods and implications of the research, the voluntariness of participation, the right to withdraw consent at any time without consequences, the funding of the research, commitments to anonymity,

[161] EU Charter, ECHR; Helsinki Declaration; European Commission, 2018, *Ethics in Social Science and Humanities*, https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020_ethics-soc-science-humanities_en.pdf.
[162] Some relevant sources are: High-Level Expert Group on AI, Ethics Guidelines for Trustworthy AI, 2019, https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai; CEPEJ, 2018, *European Ethical Charter on the use of AI in judicial systems and their environment*, https://rm.coe.int/ethical-charter-en-for-publication-4- december-2018/16808f699c; Charter of Fundamental Rights of the European Union; European Convention on Human Rights; UN Declaration of Human Rights.
[163] European Commission, 2010, *European Textbook on Ethics in Research*, pp. 11-14; Truscott, J., Graham, A. & Powell, M.A., 2019, 'Ethical Considerations in Participatory Research with Young Children', *Participatory Research with Young Children. Educating the Young Child*, Springer, 2019.
[164] European Commission, 2010, *European Textbook on Ethics in Research*, pp. 33-74.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

privacy and confidentiality of personal data, access to information, where the research findings will be published, and the name and contact details of the contact person for asking questions.[165]

As already indicated in the project proposal of ACHILLES, to obtain consent, the **following procedure** is followed to ensure voluntary consent: (1) Interested persons will be informed about the project and get details about their involvement, rights and purpose of the collection of personal data. The voluntary nature of their participation and the confidentiality of the information will be emphasised, (2) Participants are given an opportunity to ask questions, (3) A cooling-off period of at least seven days will be provided to allow participants to consider their decision, (4) Participants may ask questions again, (5) Participants provide their informed consent.

10.3 Privacy and confidentiality

Besides privacy, confidentiality and data protection being legal concerns (*infra* 5), ensuring them is also important from an ethical perspective. Legal compliance has to be strengthened and supplemented by ethical considerations. Privacy can be considered as protection over information about oneself, control over access to oneself, and control over one's ability to make important decisions about lifestyle and family. It follows that the project activities cannot result in the collection of personal data that would invade privacy. Confidentiality is closely related, and can be considered an aspect of privacy concerning the protection of personal information. Confidentiality entails the duty to keep information disclosed confidential and not reveal it to anyone. Privacy and confidentiality are important to establish and keep trust.[166] They are also linked with autonomy and human dignity: information should be provided about what will happen to personal data collected and which measures are taken to protect them.[167]

Therefore, a few measures are taken. ***Federated learning*** and ***differential privacy protocols*** will be used during the project. In M06, an early version of the ***DMP*** is drafted to set out the handling and storage of data central at all research stages. Personal data collected will be pseudonymised and encrypted before storing them, and only further processing within the project's targets will be allowed. The ***DPOs*** of partners will contribute to GDPR compliance. An ***ethical advisory board*** will be appointed that will look at (1) the design and implementation of measures to address biases), and (2) the analysis of GDPR from a more technical perspective. In the later phase of the project, discussions must be kept on the privacy and confidentiality related to the ACHILLES IDE and what kind of information will be kept,

---

[165] European Commission, 2018, *Ethics in Social Science and Humanities*, https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020_ethics-soc-science-humanities_en.pdf, 11-14.
[166] European Commission, 2010, *European Textbook on Ethics in Research*, pp. 77-93.
[167] European Commission, 2021, *Ethics and data protection*, https://ec.europa.eu/info/fundingtenders/opportunities/docs/2021-2027/horizon/guidance/ethics-and-data-protection_he_en.pdf; European Commission, 2018, *Ethics in Social Science and Humanities*, October 2018, https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020_ethics-soc-science-humanities_en.pdf, 15.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

and whether user interactions will be kept after each session or not, the way logs are kept, and the memory of the copilot assistant.

10.4 Facial recognition and verification

Specifically with regard to the identity verification use case where face verification will take place, ethical concerns may arise regarding the use of facial verification technologies. While the use case involves a one-to-one comparison of facial images with identity documents (rather than biometric identification or surveillance), it still raises questions about ***reliability, robustness, accountability and accuracy***. For instance, it has long been shown that such tools perform more accurately on white faces than on black faces due to inherent biases both in training data as in model architecture itself, or because they can pose privacy and security risks or impact human dignity and the right to integrity as well. [168] Given these risks, ACHILLES partners must take additional care to avoid them from materialising. One approach could be to focus on ***explicit consent*** (*infra* 5.1.5) while also using ***information sheets*** for participants to increase transparency and accountability about both the software and its use.

10.5 Algorithmic biases

As noted earlier (*supra* 3.5), the right to non-discrimination faces many challenges in the age of automation and AI due to the fact that AI systems can exacerbate existing stereotypes or create new ones due to biased training data or biases in model's architecture, leading to algorithmic biases, which refer to errors in the system that can produce unfair outcomes and even lead to discrimination.[169] This could result from both faulty input and model characteristics.[170] Algorithmic biases occur because (generative) AI systems typically train on vast amounts of (text) data available (on the internet) that may contain biases. As a result, AI systems' responses could inadvertently reflect and reinforce these existing biases.[171] Biased outcomes are problematic, especially in light of the trustworthiness objective of ACHILLES. Specific to the SCRIPTA use case but also the ACHILLES IDE, research has shown (*supra* 3.5) that NLP technologies can result in biases based on gender. Therefore, language diversity remains an important consideration in NLP technologies. While the availability of related tools for English has improved, other languages still lag behind due to limited data quantity and quality. These imbalances

[168] Almeida, D. Shmarko, K. & Lomas, E., 2022, 'The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: a comparative analysis of US, EU, and UK regulatory frameworks', *AI and Ethics*, pp. 377-387; Smith, M. & Miller, S., 2022, 'The ethical application of biometric facial recognition technology', *AI & Society*, pp. 167-175; Van Noorden, R. 2020, 'The ethical questions that haunt facial-recognition research', *Nature*, https://www.nature.com/articles/d41586-020-03187-3.

[169] FRA, Bias in Algorithms – Artificial Intelligence and Discrimination Report, 2022, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf

[170] Kordzadeh, N. & Ghasemaghaei, M., 2020, *Algorithmic bias: review, synthesis, and future research directions*, https://doi.org/10.1080/0960085X.2021.1927212.

[171] Wach, K., e.a., 2023, 'The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT', *Entrepreneurial Business and Economics Review* 2023, http://dx.doi.org/10.15678/EBER.2023.110201.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

for other languages result in slower, less accurate models with a higher risk of biases. However, even in English, safeguards, monitoring and regular evaluation are essential to ensure fairness and reliability.

Therefore, efforts must be made to mitigate the risk of discriminatory output due to algorithmic biases. While the data requirements in Art. 10 AI Act apply to *high-risk* AI systems (*supra* 4.1.3.2) , ensuring data accuracy and quality is important in all cases. **Data quality criteria** are: completeness, accuracy, consistency, duplication, validity (i.e. whether the data and predictions actually measure what they intend to measure), availability, and provenance. It is not just the volume of the data that matters but also its quality. While large datasets can reduce statistical uncertainty in big data applications, they do not enhance the validity of measurements if they contain significant errors in representation and measurement. In fact, poor-quality data can lead to consistently inaccurate results. To ensure high-quality training data, datasets should be properly assessed and documented, including detailed metadata and quality checks.[172] Minimising bias must be a priority during all phases of the AI system lifecycle. Additional measures to consider to ensure more transparency are scrutinising algorithms, conducting FRIAs and DPIAs, and verifying data quality (including metadata) to identify biases and abuses based on output algorithms. Providing individuals with meaningful information about AI outputs can enhance accountability and fairness as well.[173]

10.6 Hallucinations

Hallucinations are a specific risk in relation to generative AI where the **AI system generates inaccurate or fictional but highly realistic output**, such as names, dates, and events.[174] Within ACHILLES, this issue is relevant to different activities, such as the SCRIPTA and HERA use cases and the ACHILLES IDE, as in these cases, the LLMs underlying the systems could generate highly realistic output but incorrect – which could undermine their overall trustworthiness, and the accuracy and reliability of their output.[175]

To mitigate the risk of hallucinations, different measures could be considered. These could include ensuring that training comes from high-quality and reliable sources, conducting regular check on the IDE's underlying models, and implementing a human-in-the-loop approach for validation. Other

[172] FRA, 2018, *#BigData: Discrimination in data-supported decision* making, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-focus-big-data_en.pdf; FRA, Bias in Algorithms – Artificial Intelligence and Discrimination Report, 2022, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2022-bias-in-algorithms_en.pdf

[173] FRA, 2018, *#BigData: Discrimination in data-supported decision making*, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-focus-big-data_en.pdf.

[174] Keary, T., 2024, 'Hallucinations (Artificial Intelligence)', *Techopedia*, https://www.techopedia.com/definition/ai-hallucination#:~:text=Generative%20AI-driven%20chatbots%20can%20fabricate%20any%20factual%20information%2C,produce%20inaccurate%20information%2.

[175] Feldman, P., e.a., 2023, *Trapping LLM Hallucinations Using Tagged Context Prompts*, www.researchgate.net/publication/371490092_Trapping_LLM_Hallucinations_Using_Tagged_Context_Prompts/citations.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

measures to be considered are integrating contextual prompts, constraints and custom embeddings, or RAG approaches. The technical partners can further evaluate the system for logical coherence while designing it to cite its sources when generating responses.

10.7 Trustworthiness

One of ACHILLES' objectives is to ensure a rich ecosystem for trustworthy AI and to provide a technically and socially robust environment, i.e. free of cyber-attaches and biases and striving for transparency of AI-based systems. Hence, for the design, development and use of AI systems in ACHILLES to be ethical, the trustworthiness of the AI system plays a crucial role.

In this regard, the **Ethics Guidelines for Trustworthy AI** are most relevant, as they provide a framework for developing AI systems in a trustworthy manner – and on which the AI Act is actually based.[176] The Guidelines were issued by the High-Level Expert Group on Artificial Intelligence (AI HLEG), a group of 52 experts mandated to make recommendations on the regulation and governance of AI.[177] In order to be deemed trustworthy, AI systems must (1) be lawful and respect all applicable laws and regulations; (2) be ethical and respect all ethical principles and values; and (3) be robust from a technical perspective and consider the social environment. The Guidelines build upon fundamental rights, such as respect for human dignity, freedom of the individual, respect for democracy, justice and the rule of law, equality, non-discrimination and solidarity, and citizens' rights. They also emphasise the importance of the four ethical principles, rooted in fundamental rights, which must be protected to ensure that AI systems are developed, deployed and used in a trustworthy manner. These are respect for human autonomy, prevention of harm, fairness and explicability. Based on these principles, the AI HLEG has developed seven key requirements that AI systems must meet to be trustworthy. These requirements will guide the design and development of AI systems throughout the ACHILLES project, such as the IDE and use cases. To ensure that these principles are operationalised in a practical and consistent manner, the AI HLEG also developed the Assessment List for Trustworthy AI (ALTAI). ALTAI provides a structured self-assessment tool to evaluate how the seven key requirements for Trustworthy AI are being met in real-world applications. As elaborated further in D4.4, ACHILLES will use ALTAI as a reference framework to support the ethical dimensions of system development, validation, and pilot deployment.

Please note that for the use case validation (Task 7.4), the *Z-Inspection Process for Trustworthy AI* will be carried out by ARCADA to ensure the use cases are trustworthy and align with EU values and

[176] Smuha, N.A, 2019, 'The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence', *Computer Law Review International*, pp. 97-106; Smuha, N.A., e.a., 2021, *How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991
[177] Renda, A., 2024, 'Europe: Toward a Policy Framework for Trustworthy AI', *The Oxford Handbook on Ethics of AI*. OUP.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

definitions of trustworthy AI. However, since the AI HLEG principles have a broader scope, the focus in the following sub-sections lies on those principles.

10.7.1 Human agency and oversight

The principle of respect for human autonomy requires that AI systems should support **human autonomy** and decision-making.[178] Users should be allowed to make *informed* autonomous decisions regarding the AI system at issue. Moreover, human oversight should be provided through governance mechanisms such as a *human-in-the-loop, human-on-the-loop or human-in-command approach*. To ensure human agency, transparency about the AI systems is crucial and adequate information must be provided to users. It also follows that the technical partners should apply a human-in-the-loop approach and do regular *checks* of their systems to fine-tune and further enhance their performance.

10.7.2 Technical robustness and safety

Closely linked to the principle of prevention of harm is the requirement for technical robustness and safety. This requirement refers to **resilient and secure** AI systems that are safe, accurate, reliable and reproducible. AI systems should be protected against vulnerabilities, e.g. hacking or data leakage. Appropriate safety safeguards, e.g. a fallback plan, must be put in place. The technical and safety safeguards of the technical partners to develop a resilient and secure AI system are hence important.

10.7.3 Privacy and data governance

AI systems must guarantee privacy and data protection throughout the entire life cycle of the system (*supra* 3.4 and 5). In addition, **adequate data governance** mechanisms must be deployed to ensure the quality and integrity of data. To ensure legitimised access to data, *data protocols* governing data access should be put in place that clearly define who has access to what data. In addition to the DMP (which will be updated throughout the project), partners should have data governance measures in place to ensure adequate privacy and data governance. As explained, the *quality and accuracy of the data* used for developing AI systems must be sufficiently considered, so that they contain as few biases or inaccuracies as possible.

10.7.4 Transparency

To meet the principle of **explicability**, data, system, and AI business models must be transparent. Transparency means that data sets should be *documented* to allow for traceability. Furthermore, the technical processes of AI systems should be explainable to human beings. Humans also have a right to be *informed* that they are interacting with an AI system and the characteristics of that AI system. With

[178] Calvo, R.A., Peters, D., Vold, K. & Ryan, R.M., 2020, 'Supporting Human Autonomy in AI Systems: A Framework for Ethical Enquiry', *Ethics of Digital Well-Being. Philosophical Studies Series*, Springer, https://doi.org/10.1007/978-3-030-50585-1_2.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

regard to explainability and communication, the project will always try to inform users adequately about the AI system and involve them where possible. It has indeed been recognised that (algorithmic) transparency is essential for enabling citizens' oversight over designs, as it allows the exercise of fundamental rights (e.g. non-discrimination because you can scrutinise the AI system), and it enhances accountability because reasons must now be given for decisions.[179]

Therefore, participants involved in use case validation should be informed about the system's capabilities and limitations so as not to raise unwarranted expectations.

Explainability is one important element of transparency. It concerns the ability to explain the technical processes of an AI system and related human decisions. The research line, known as Explainable AI (XAI), revolves around the idea that if an AI system has significant influence over human lives, users must be able to understand and request explanations on its decision-making process. XAI aims to provide interpretability and explainability through textual and visual cues.

To support this, WP5, "*Explanations, Evaluation and Reporting*", is dedicated to developing, implementing, and evaluating XAI methods in the project to enhance trust, provide human-readable explanations and improve data documentation through reporting. The ACHILLES project will bridge the gap between *local and global XAI* by deriving global concepts from local explanations. These techniques promote trustworthiness, confidence, fairness, and help identify dataset issues, increase model efficiency, and improve model reasoning.

One of the research questions that will also be further explored in D4.3, i.e. the doctrinal research, is the identification of to what extent ML systems must explain their reasonings to meet principles of transparency and accountability. It will also explore what explainable AI should look like within the context of ACHILLES, the types of explanations required for the transparency principle, and the specific circumstances, scenarios, and target audience that determine the nature of these explanations.[180]

10.7.5 Diversity, non-discrimination and fairness

Inclusion, diversity, non-discrimination and fairness should be integrated into the AI system's life cycle at all times. This means that *unfair bias must be avoided in the data sets and the development* of the AI system. Measures must be taken to ensure that *data sources* contain as few biases as possible to avoid discriminatory outcomes. As mentioned, much attention in ACHILLES is paid to ensuring **bias detection and mitigation** (*supra* 3.5) with an entire dedicated task 1.2 for bias detection and mitigation. One of the research questions that will also be further explored in D4.3, i.e. the doctrinal research, is exploring the relevant criteria for detecting biases in systems' datasets to ensure fair AI decision-

[179] GPAI, 2024, *Algorithmic Transparency in the Public Sector A state-of-the-art report of algorithmic transparency instruments*, https://wp.oecd.ai/app/uploads/2024/12/14-Algorithmic-Transparency-in-the-Public-Sector-A-state-of-the-art-report-of-algorithmic-transparency-instruments.pdf.
[180] Hacker, P. & Passoth, J.H., 2020, 'Varieties of AI Explanations under the Law. From the GDPR to the AIA, and Beyond', *Lecture Notes on Artificial Intelligence 13200: beyond explainable AI*, Springer.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

making and prevent discrimination. In addition, the project will explore privacy-preserving techniques to guarantee the anonymity of personal data and ensure compliance with GDPR.

Diversity entails employing people from *diverse backgrounds*, cultures and disciplines when developing an AI system. Therefore, ACHILLES consists of a collaboration of research centres (FhAICOS, FhHHI, INESC-ID), academic institutions (KUL, UDS, ISRUC, ARCADA), industrial partners (IDNOW), associations (SERMAS), tech/SMEs (AXIOLOGIC, ETICAS, CTTI, INNEN, PNO, CUOMOIT) and a trusted certification authority as associated partner (INCM). This diversity must ensure a robust common platform and a combination of theoretical and practical expertise. Diversity also requires AI systems to be accessible to all, regardless of any disability, and relevant stakeholders must be involved throughout their existence.

As an overarching principle, **fairness** involves equality and equity, emphasising that all individuals must be treated *equally*[181] – which links to the importance of detecting biases and preventing non-discrimination. The substantive notion of fairness requires that the content and outcomes of decisions are equitable and not based on subjective choices. However, it is important to note that in designing and developing AI systems, this is rather an ideal that is difficult to achieve since all systems embed values, whether or not intentional.[182] To address this, Task 4.4 and the corresponding Deliverable 4.6 focus on human-centred and value-sensitive design to ensure values behind decisions affecting humans are accounted for. The growing field of value-sensitive design (VSD) – exploring how human values, needs, preferences and cultural backgrounds may shape a system's design – will be explored in greater detail in its own Deliverable.

10.7.6 Societal and environmental well-being

The broader society and the environment should be considered as well throughout the AI system's life cycle. AI systems should benefit all human beings while being **sustainable and environmentally** friendly. This links to the UNESCO recommendations (*supra* 4.2.1) that also mention the importance of the environment and ecosystems in the context of AI development, e.g. due to energy demands and power.[183] They call for actions, such as assessing the (in)direct impact throughout the AI system's lifecycle. In line with this principle, the project aims to develop AI systems that are energy efficient, and whose energy and resource consumption is limited.

10.7.7 Accountability

[181] Kitting, M., e.a., 2024, 'Assessing trustworthy AI: Technical and legal perspectives of fairness in AI', *Computer Law & Security Review*.
[182] Naudts, L. & Vedder, A., 2024, 'Fairness and Artificial Intelligence', *The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence*, pp. 79-100.
[183] UNESCO, 2023, *The UNESCO Recommendation on The Ethics of AI: Shaping the Future of Our Societies*, https://www.unesco.nl/sites/default/files/inline-files/Unesco%20AI%20Brochure.pdf.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Finally, the requirement of accountability demands that mechanisms be implemented to ensure ***responsibility*** and accountability for AI systems and their outcomes. *Auditability* requires the ability to assess algorithms, data and design processes. Various systems can be established to guarantee accountability. Partners could consider conducting *internal or external audits*, *reports* or impact *assessments*. The AI systems must be regularly evaluated throughout the development phase. Automated reporting can help address challenges and facilitate transparent communication between AI systems developers and users. In addition, the ethical AI impact assessments, conducted in D4.4, will ensure accountability within ACHILLES.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

11 CONCLUSIONS

As conclusions, a concise table will be provided which indicates which legislation is likely applicable and what requirements and obligations apply to the project activities.

Fundamental rights	Fundamental rights must be prioritised in all phases of the ACHILLES project, from the design, development, deployment and use of AI systems and beyond. Special attention should be given to the following fundamental rights – especially in light of the context of automation: ▪ Human dignity and the right to the integrity of the person require that all project activities but also the development of AI systems (both the ACHILLES IDE, use cases and AI systems developed through the IDE) respect human dignity and the integrity of the person. Specifically, in the context of automation, a human-centred approach should be taken where individuals remain at the centre of all AI-related discussions. Individuals cannot be made subject to AI systems without their knowledge or informed consent. ▪ Privacy, family life and personal data protection require that privacy and family life should always be respected throughout all project activities. It requires the fair processing of personal data, subject to multiple processing principles (as more elaborately explained under the GDPR section hereunder). ▪ Equality and non-discrimination require that throughout ACHILLES – whether it be the IDE itself, the AI systems developed through the IDE or all other project activities – individuals must be treated equally and fairly and not subject to direct or indirect discrimination. Specifically for the context of automation, discriminatory outcomes in AI systems – due to data bias or model bias – must be mitigated and avoided. Special attention must be paid to bias detection and mitigation through different techniques, such as data quality and accuracy, adversarial training or continuous testing. ▪ Intellectual property requires that everyone should enjoy their possessions, including literacy and artistic property. AI systems cannot infringe IP rights. ▪ Healthcare requires that everyone has the right to access to healthcare and benefit from medical treatment. For the context of automation, medical AI systems should not compromise

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	human autonomy, transparency, explainability and non-discrimination. ▪ Environmental protection requires – especially in context of AI systems – that AI development should be sustainable and avoid overly ecological impact. Sustainable practices adoption is therefore important for ACHILLES for IDE and AI systems developed through IDE.
Artificial Intelligence	Since ACHILLES' main purpose is to develop an IDE that can provide information to AI system developers, both the IDE and the AI systems developed through the IDE should comply with AI requirements. EU AI Act The EU AI Act imposes different obligations on AI systems. As a first step, it should be assessed whether an AI system falls under the territorial and personal scope of the Act. If falling under the AI Act's scope, the obligations differ based on the risk the AI system poses to health, safety and fundamental rights. AI systems with unacceptable risk are prohibited, high-risk AI systems are subject to mandatory requirements and ex-ante conformity assessments, limited-risk AI systems must comply with transparency obligations, and minimal-risk AI systems can voluntarily choose to comply. In addition, GPAI models and systems are subject to an additional set of obligations. ▪ For each AI system, a provider should be identified since they have to comply with most of the provisions of the AI Act. Providers are the ones developing the AI system or GPAI model, or they have it developed by third parties but place it on the market. ▪ For high-risk AI systems, different obligations apply. ○ A risk management system must be established, implemented, documented and maintained with regard high-risk AI systems. This is a continuous process to identify and analyse the risks of the system and to identify mitigation measures to target risks (Art. 9). ○ Data and data governance requirements require that when systems are developed on the basis of training, testing data sets must meet quality criteria and be subject to data governance and management practices. It also requires that the training, validation and testing are

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	relevant, sufficiently representative and as much as possible free from errors and complete (Art. 10). ○ Technical documentation and record-keeping (automatic logs) should be kept to verify compliance (Art. 11, 12, 17, 18, 19). ○ Transparency and information obligations towards deployers require AI systems must be designed and developed in such as way that operations are sufficiently transparent to enable deployers to interpret system's output and use it appropriately. Information should allow deployers to understand the system's working, functions and limitations (Art. 13). ○ Human oversight requires that AI systems must be designed and developed to that they can be overseen by natural persons. Human oversight measures can therefore be adopted, such as in-built operational constraints or ensuring competence of human overseers (Art. 14). ○ Accuracy, robustness and cybersecurity obligations require that AI systems must be designed and developed so that they are accurate, robust and cybersecure. Systems must be resilient against errors or inconsistencies. Technical and organisational measures must be taken, such as backup or fail-safe plans (Art. 15). ○ A quality management systems must be put to ensure compliance with the AI Act, including a strategy and action points for compliance (Art. 17). ○ Besides providers, deployers of high-risk AI systems are subject to certain obligations. They must monitor the operation of systems based on instructions for use and therefore take technical and organisational measures. They must assign human overseers. Under certain circumstances deployers will be considered providers and subsequently all obligations will be transposed to them (Art. 26). ○ Deployers of high-risk AI systems must carry out a fundamental rights impact assessment (FRIA), i.e. an assessment to identify specific risks to fundamental rights and appropriate measures. This obligation is limited to deployers that are bodies governed by public

	law or private entities providing public services, and deployers of high-risk AI systems for evaluating creditworthiness, establishing credit scores or for health insurance risk assessment and pricing (Art. 27). ▪ For certain AI systems, transparency obligations apply. AI systems intended to directly interact with natural persons must be designed to inform users that they are interacting with AI system rather than a human (Art. 50). ▪ For all AI models and systems, efforts should be paid to ensure energy sustainable AI systems and reduce resources, energy consumption, and develop as energy-efficient models and systems as possible. International regulatory initiatives International regulatory initiatives complement the EU AI Act with regard to the regulation of AI systems. While they might be non-binding, attention should still be paid given their importance at international level. Standards Standards can play an important role in regulating AI systems as they provide more information about implementation of technical aspects. ISO/IEC 42001 on AI Management Systems, for instance, provide information about managing AI projects and AI risk assessments. At EU level, CEN-CENELEC is currently drafting standards to implement technical aspects of EU AI Act, but they are not yet published.
Privacy and data protection	The GDPR sets out the rules for processing personal data and the persons responsible for ensuring compliance with the regulation. It does not apply to anonymous data. Both 'processing' and 'personal data' are interpreted broadly. For each processing activity, data controller(s) and potential processor(s) should be identified, as they are responsible for ensuring compliance with the GDPR. Data controllers decide the *purpose* and *means* of data processing (i.e. cumulative conditions). If two or more entities decide autonomously the purposes and means

	of data processing, they are separate data controllers. If determined together, they are *joint controllers* and must set out their respective responsibilities. The following principles must always be respected when engaging in processing activities (Art. 5): ▪ Lawfulness, fairness and transparency: an appropriate legal basis must be determined before each processing operation, whereby the interests of the data subjects must be taken into account. Transparency requires data controllers to inform data subjects about processing activities and their rights. ○ In case consent is chosen as legal basis for the processing, consent must meet the following conditions (Art. 6): ▪ Consent must be a clear affirmative act ▪ Freely given, having real choice to consent ▪ Specific, for specific purpose ▪ Informed, about purposes and activities in clear and plain language information ▪ Unambiguous indication of agreement, no doubt ▪ Consent can be withdrawn at any moment, as easy as giving consent. ▪ Consent within the GDPR is not be confused with the notion of informed consent as an ethical requirement for human participants in research activities. ▪ Purpose limitation: personal data can only be processed for a specific, explicit and legitimate purpose. ○ The further processing for other purposes that are not compatible with the original purpose is in principle prohibited. There are two exceptions, namely when further processing is based on consent, or when Union or Member States law constitutes a necessary and proportionate measure in democratic society. If the two exceptions do not apply, then the controller must carry out a compatibility assessment (Art. 6(4)). For the further processing in scientific research, more lenient rules apply (Art. 89). ▪ Data minimisation: processing of personal data must be limited to what is necessary to achieve the specific purpose(s). The

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	period for which personal data is stored must be limited to the minimum. Different techniques can be used, such as anonymisation. ▪ Accuracy: data controllers must ensure that personal data is kept up to date and accurate, and conduct regularly checks. ▪ Storage limitation: personal data should only be stored for the period of time that is necessary for the processing purposes. Afterwards, they must be deleted or anonymised. ▪ Integrity and confidentiality: measures must be taken to ensure security of personal data against the unauthorised or unlawful processing, accidental loss, destruction or damage. ▪ Data protection by design and by default: data controllers must ensure principles already in the design of any processing activity. This implies that technical and organisational measures must be implemented, such as pseudonymisation, data minimisation, encryption, when determining means of processing and processing itself. The following rights towards data subjects must be guaranteed. ▪ Right to transparency and information: data subjects must receive information related to processing of personal data in a transparent, concise, intelligible and easily accessible manner, using clear and plain language. Information can be complemented by information sheets (Art. 12, 13, 14). ▪ Right of access: data subjects must have access to personal data so to exercise their rights (Art. 15). ▪ Right to rectification: data subjects have a right to rectify inaccurate data from data controllers (Art. 16). ▪ Right to erasure: data subjects have a right to get their personal data erased when no longer necessary for processing purpose, consent has been withdrawn, or processing was unlawful (Art. 17). ▪ Right to restriction: processing of personal data can be restricted in certain circumstances, e.g. when data subject contests accuracy personal data, processing is unlawful etc. (Art. 18). ▪ Notification obligation: data controllers must communicate any rectification or erasure of personal data or restrictions of processing to data subjects (Art. 19). ▪ Right to data portability: data subjects can ask personal data and transmit it to another data controller when the processing is

	based on consent and is carried out based through automated means (Art. 20). ▪ Right to object: data subjects can object to processing when it is based on the legal basis of public interest, or for lelegitimate interests or direct marketing purposes. ▪ There is a general prohibition on fully automated decision-making (Art. 22). Special categories of personal data Special categories of personal data, such as biometric and health related data, can in principle not be processed, unless under a limited number of conditions, such as explicit consent of data subjects, i.e. express statement of consent, ideally in written format (Art. 9). Synthetic data Whether synthetic data is considered pseudonymised (and thus falling under GDPR), or anonymised (and not subject to GDPR) depends on the circumstances. When synthetic data has no direct one-to-one link to individuals, it can be considered anonymous, because it is indistinguishable from the original data. However, others have argued that it should not be considered anonymous data because one-to-one relationships are still possible to derive when synthetic data set still has characteristics of original data with high accuracy. Hence, whether synthetic data is anonymised or pseudonymised depends on how much it deviates from the original data, with the key factor being whether identifiability if effectively prevented and anonymity is sustained over time. Data protection officer A DPO has to be appointed by the data controller or processor. A DPO is obliged to oversee the company's data activities. It is recommended to appoint a DPOs when large-scale processing of special categories of data under Art. 9 are concerned, such as health related or biometric data. Data protection impact assessment A DPIA is a systematic description of envisaged processing activities and purposes. It is required under certain conditions, namely in case

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	of systematic and extensive evaluation of personal data based on automated processing, processing on a large scale of special categories of data, or systematic monitoring of publicly accessible area on a large scale (Art. 35).
Data governance	The Data Act and Data Governance Act facilitate reliable and secure access to data, which is relevant for the development of AI systems. Data Act The DA regulates data access and use, and concerns data generated using a product or related service. Different obligations and requirements apply. ▪ Chapter II includes provisions on the *rights of users* to use data connected products and related services. o Connected products and related service data must be designed and manufactured to ensure that users can access them easily, securely, and free of charge, in a comprehensive, structured, commonly used and machine-readable format (Art. 3). o Data holders must make data available to users when data cannot be directly accessed by them from the connected product or related service (Art. 4). o Users can share data with third parties under certain conditions (Art. 5). o Third parties can only process the data made available to them (pursuant to Art. 5) for the purposes and conditions agreed with the user and in accordance with law (Art. 6). ▪ Chapter III includes horizontal obligations for data holders to make data available in *business-to-business* relations. o Data holders must make data available to data recipients when they are obliged to do so according to Art. 5. Arrangements must be agreed upon, but they should be under fair, reasonable and non-discriminatory terms and conditions and in a transparent manner, without discrimination, and under no exclusive basis or without IP infringements (Art. 8). o Compensation for making data available must be non-discriminatory and reasonable (Art. 9). o Data holders may apply technical protection measures, such as smart contracts and encryption, to prevent

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	unauthorised access to data and to ensure compliance with obligations (Art. 11). ▪ Chapter IV deals with the *unfair contractual terms* related to data access and use between enterprises. Such terms are non-binding on disadvantaged enterprise (Art. 13). ▪ Chapter V includes obligations to *make data available to public sector bodies*, the EC, European Central Bank and Union bodies on exception need. o Conditions are set out for exceptional need, such as a public emergency (Art. 14-18). o The bodies cannot use data for other purposes then requested, and must implement technical and organisational measures to preserve confidentiality and integrity data, erase data when no longer necessary for purpose (Art. 19). o Data holders are entitled to fair compensation to cover technical and organisational costs (Art. 20). ▪ Chapter VI foresees the rules for *switching* between data processing services (Art. 23-31). ▪ Chapter VII includes rules related to *unlawful international governmental access and transfer* of non-personal data (Art. 32). ▪ Chapter VIII includes the provisions on *interoperability* of data, data sharing mechanisms and services and common European data spaces. Essential requirements regarding interoperability of data and data sharing mechanisms and services are set out (Art. 33). These concern: o the dataset content use restrictions, licences, data collection methodology, data quality and uncertainty shall be sufficiently described to allow the recipient to find, access and use the data, o the data structures, data formats, vocabularies, classification schemes, taxonomies and code lists must be described in a publicly available and consistent manner, o the technical means to access the data, such as application programming interfaces, and their terms of use and quality of service shall be sufficiently described to enable automatic access and transmission of data between parties, including continuously, in bulk download or in real-time in a machine-readable format

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	where that is technically feasible and does not hamper the good functioning of the connected product, and ○ the means to enable the interoperability of tools for automating the execution of data sharing agreements, such as smart contracts shall be provided <u>Data Governance Act</u> The DSA regulates processes and structures to facilitate voluntary data sharing by individuals and businesses. ▪ Chapter II entails the provisions on the *re-use* of certain categories of protected data by public sector bodies. ○ The chapter only applies to data held by public sector bodies which are protected on grounds of (a) commercial confidentiality, including business, professional and company secrets; (b) statistical confidentiality; (c) the protection of intellectual property rights of third parties; or (d) the protection of personal data. ○ Public sector bodies must make publicly available the conditions for re-use and procedures to request re-use of data. These conditions must be non-discriminatory, transparent, proportionate and objectively justified in light of categories of data, purposes of re-use and nature of the data for which re-use is allowed. The protected nature of data must always be protected, e.g. by using anonymisation or within a secure processing environment. Re-users are prohibited to re-identify any data subjects to whom the data relates, and must take technical and operational measures to prevent re-identification and to notify any data breach resulting in the re-identification of the data subjects concerned to the public sector body. Re-use must also be in compliance with intellectual property rights. Confidential data cannot be disclosed as a result of allowing re-use. If public sector body cannot grant access to certain data to re-use, they should help seek consent to re-use data (Art. 5). ○ Public sector bodies can charge fees to re-use data (Art. 6). ▪ Chapter III includes the requirements applicable to *data intermediation services*. They must act as trustworthy neutral

	third party to connect individuals and companies with data users (Art. 10-11). ▪ Chapter IV includes the provisions related to *data altruism*, which refers to individuals and companies giving consent or permission to make available data that they generate (voluntarily and without reward) to be used for objectives of general interest (Art. 17-19). European Common Data Spaces Within the European strategy for data, the EU is trying to establish Common European Data Spaces, which aims to create an internal market for data, making more data available for access and re-use in a trustworthy and secure environment for the benefit of European businesses and citizens. The common data spaces are meant to be a digital environment, open for all, having a secure and privacy-preserving infrastructure to pool, asses, share, process and use data. They should have fair, transparent, proportionate and non-discriminatory access rules, enabling data holders to grant access and share certain (non-)personal data.
Information society services	The Digital Services Act and Digital Markets Acts regulate information society services. They aim to create a secure digital space that protects users' and businesses' fundamental rights while promoting innovation. They include provisions on facilitating access to data, which can be important for designing, validating and training AI models and systems. Digital Services Act The DSA regulates online intermediaries and platforms. It aims to prevent illegal activities online and ensure fundamental rights online. Depending on the role, size, impact of online services, different obligations apply. ▪ Intermediary services: o Providers have to designate a single point of contact to communicate with authorities (Art. 11), a single point of contact for recipients of the service (Art. 12), and a legal representative in case they are not established in the EU (Art. 13).

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	○ They must include information on any restrictions in relation to the use of their services to recipients in their terms and conditions (Art. 14). ○ They are subject to transparency reporting obligations in the form of making publicly available a clear and comprehensible report of any content moderation that they did. ▪ Hosting services ○ Providers must put in place mechanisms to allow individuals to notify them of illegal content (Art. 16). ○ They must provide information to individuals when there are any restrictions imposed on the information provided by the recipients when it is illegal or incompatible with their terms and conditions (Art. 17). ○ They must inform the law enforcement authorities when they become aware of information giving rise to suspicion that a criminal offence involving a threat to the life or safety to an individual has taken place (Art. 18). ▪ Online platforms ○ Providers must provide recipients of the service access to an effective complaint-handling system to lodge complaints against decisions taken by them, such as decisions on removing or disabling access or visibility of information, termination of services or recipient's account, about monetisation information (Art. 20). ○ Recipients have the right to opt for out-of-court dispute settlement bodies to resolve disputes relating to those decisions (Art. 21). ○ Providers must take necessary technical and organisational measures to ensure notices submitted by trusted flaggers are given priority (Art. 22). ○ Providers have a right to suspend services to recipients who frequently provide manifestly illegal content (Art. 23). ○ Providers must abstain from designing online interfaces that deceives or manipulates recipients of services (Art. 25). ○ They must ensure that advertisements on their online interfaces are clearly presented as advertisements (Art. 26).

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	○ In case providers make use of recommender systems, they must set out in terms and conditions, main parameters used in this system and any options for recipients to modify or influence these (Art. 27). ○ For children, providers must put in place measures to ensure high level of privacy, safety and security on their service (Art. 28). ▪ Very large online platforms and very large online search engines ○ Providers of VLOPs and VLOSEs are subject to carrying out risk assessment (Art. 32) and take appropriate mitigation measures (Art. 33). ○ They must also allow independent audits to assess compliance with obligations (Art. 37). ○ To ensure efficient audits, platforms/engines should provide access to all relevant data (Art. 40). <u>Digital Markets Act</u> The DMA regulates gatekeepers, i.e. undertakings offering core platform services and being designated by the EC. They are subject to obligations, such as allowing third parties to inter-operate with the gatekeepers' own services, or allow their business users to promote their offer and conclude contracts with their customers outside the gatekeeper's platform. In addition, they are subject to prohibitions, such as treating their own products more favourably in ranking than third parties' or prevent users from un-installing any pre-installed software. In the context of ACHILLES, it seems unlikely that the ACHILLES IDE will qualify as a gatekeeper given the strict criteria for it to apply.
Cybersecurity	In addition to Art. 15 AI Act, which imposes cybersecurity obligations on AI systems, specific cybersecurity legislative frameworks apply to the ACHILLES project, namely the NIS2 Directive, the EU Cybersecurity Act and the EU Cyber Resilience Act. <u>NIS2</u> NIS2 imposes cybersecurity requirements for entities that operate in critical and important sectors, lays down cybersecurity risk-

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	management measures and reporting obligations for certain entities, and obligations on cybersecurity information sharing. NIS2's scope of application is limited to entities that operate in one of the critical sectors included in the Annexes, such as energy, transport, banking, financial market infrastructure, health, drinking water, waste water, digital infrastructure, ICT service management, public administration, and space, as well as postal and courier services, waste management, chemicals, food, manufacturing, digital providers, research. In that case, several obligations apply. Entities must: ▪ implement a comprehensive proactive risk management, meaning to implemented industry and adequate cybersecurity measures (Art. 21). ▪ report any incident that has significant impact on the provisions of their services: they should issue without undue delay (within 24h) an early warning, an incident notification (within 72h), issue upon request an intermediate report on status of incident and crisis management ▪ submit a final report on incited, mitigation measures and effects (Art. 23). <u>EU Cybersecurity Act</u> The EU Cybersecurity Act establishes a new and stronger mandate for the EU agency for cybersecurity, i.e. ENISA, and introduced a framework for voluntary European cybersecurity certification schemes for ICT products, services and processes, yet does not offer concrete cybersecurity requirements relevant to ACHILLES. <u>EU Cyber Resilience Act</u> The Cyber Resilience Act imposes conditions for the development of secure hardware and software ICT products to ensure that manufacturers take security measures throughout product's lifecycle, and conditions to allow users to take cybersecurity into account when selecting and using digital products.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	A product with digital elements is defined as "*software or hardware product and its remote data processing solutions, including software or hardware components being placed on the market separately.*" These products can only be made available on the market when they meet a set of cybersecurity requirements, such as: ▪ have a secure default configuration, without known exploitable vulnerabilities, mitigation measures for vulnerabilities, control mechanisms for unauthorised access, mechanisms to ensure confidentiality, integrity and availability and so on (Art. 6 j Part I of Annex I). ▪ have processes put in place by the manufacturer to comply with the vulnerability handling requirements set out in Part II of Annex I that focus on identifying and documenting vulnerabilities and risks, applying regular tests and reviews of security, taking measures to facilitate sharing information about vulnerabilities, and ensuring security updates. ▪ Some products are subject to conformity assessment procedures (Art. 7(1) j Annex III j Art. 32). These include password managers, software searching for malicious software, VPN, network management systems, boot managers, operating systems, smart home general purpose virtual assistants, identify management systems and privileged access management software, including authentication and access control readers, including biometric readers, digital certificate issuance software, physical or virtual network interface and so on. ▪ In case the products fulfil essential cybersecurity requirements in Part I of Annex I, processes put in place by manufacturer comply with essential cybersecurity requirements in part II of Annex II, and there is a EU declaration of conformity issued under Cyber Resilience Act, it is deemed that product complies with cybersecurity requirements set out in Art. 15 AI Act.
Sectoral legislation	The AI systems being developed and the use cases within ACHILLES may trigger sectoral laws in addition to EU horizontal regulations. Although it must be further discussed and assessed which sectoral legislation can be relevant and to what extent compliance checks should go in the context of sectoral legislation, it is already helpful to detect some of the key sectoral laws that may apply.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	<u>Medical Devices Regulation</u> The MDR may be relevant for use cases concerning healthcare and pharmaceutics. The MDR lays down rules for placing on the market, making available on the market or putting into service medical devices for human use and accessories for such devices in the EU. Medical devices that incorporate AI are classified as high-risk and must comply with strict regulatory requirements to ensure their safety and effectiveness. Obligations towards manufacturers of medical devices range from establishing a risk management, keeping technical documents and EU declaration of conformity, drawing up a quality management system, reporting incidents, conducting a conformity assessment and so on. <u>Copyright legislation</u> Although copyright legislation is primarily regulated at national level, several EU directives aim to harmonise these rules across EU Member States. Copyright legislation can be relevant for the SCRIPTA and HERA use cases and the ACHILLES IDE, as well as for other AI systems being developed through the IDE in the future. Several EU *copyright* directives exist. The Directive on Copyright and Related Rights in the Digital Single Market includes obligations relevant to the context of AI, such as when the exception for text and data mining applies that allows training of AI systems on IP protected works, measures relating to certain uses of protected content by online services, and a new obligation towards online content-sharing service providers. The latter must obtain authorisation of rightsholder when they allow public to access works that are protected by copyright and uploaded by their users. For *trade secrets*, the Directive 2016/943 offers protection against unlawful acquisition, use and disclosure of trade secrets. And for *databases*, the Directive 96/9/EC establishes a sui generis right that protects the content of the database, including computer programs, safeguarding them against unauthorised use.
Ethics considerations	In addition to legal obligations, ethical principles have to be respected, such as privacy, confidentiality, security, integrity, respect for human dignity, quality, non-discrimination, diversity, human autonomy, human agency and oversight, inclusiveness, harm

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

	prevention, fairness, accountability, transparency, accountability and so on. *Ethical consent* is required for human participants involved in the use cases validation – in addition to the informed (or explicit) consent required for the processing of personal data according to the GDPR. A valid ethical consent requires adequate information, must be voluntary and free, and participants must have sufficient competence to consent. *Privacy and confidentiality* are to be protected through different measures, such as federated learning, differential privacy protocols, pseudonymisation and anonymisation, encryption and so on. While the identity verification use case does not involve facial recognition in the strict sense (i.e. retrieving an identity from a database based on facial images), it does rely on *facial verification technologies*, referring to one-to-one matching between a facial image or video and an identity document. In such contexts, additional care must be taken to avoid risks relating to reliability, robustness, accountability and accuracy – such as using explicit consent and information sheets, and ensuring data quality. *Algorithmic biases* can arise both from biases in data and in the architecture of algorithms. Biases can be mitigated through different measures, such as ensuring data quality, accuracy and representativeness, as well as implementing techniques such as data augmentation (e.g. using synthetic data to balance underrepresented groups) and model-based techniques (e.g. adversarial debiasing or fairness-aware optimisation). Specific to generative AI, measures must be implemented to mitigate the risks arising from *hallucinations*, such as data quality, regular checks of the underlying model, human-in-the-loop approaches, contextual prompts, constraints, custom embeddings, systematic checks or quoting sources. Lastly, given ACHILLES' objective to contribute to the *trustworthy* AI debate, the Ethics Guidelines for Trustworthy AI issued by the High Level Expert Group on Artificial Intelligence must be respected.

Funded by the European Union
This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Table 3 – Conclusions

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

12 REFERENCES

Legislation

International

Charter of Fundamental Rights of the European Union, OJ C 326, 26.10.2012, p. 391–407.

Council of Europe, 2018, Convention 108 + Convention for the protection of individuals with regard to the processing of personal data, https://rm.coe.int/convention-108-convention-for-the-protection-of-individuals-with-regar/16808b36f1.

Council of Europe, 1950, *European Convention on Human Rights*. https://www.echr.coe.int/documents/d/echr/Convention_ENG.

Council of Europe, 2024, *The Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law*, https://rm.coe.int/1680afae3c.

UNESCO, 2022, *Recommendation on the Ethics of Artificial Intelligence*, https://unesdoc.unesco.org/ark:/48223/pf0000381137.

OECD, 2019, *AI Principles*, https://www.oecd.org/en/topics/ai-principles.html.

OHCHR, 1966, International Covenant on Civil and Political Rights. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-civil-and-political-rights.

OHCHR, 1966, International Covenant on Economic, Social and Cultural Rights. https://www.ohchr.org/en/instruments-mechanisms/instruments/international-covenant-economic-social-and-cultural-rights.

OHCHR, 1966, Optional Protocol to the International Covenant on Civil and Political Rights, https://www.ohchr.org/en/instruments-mechanisms/instruments/optional-protocol-international-covenant-civil-and-political.

OHCHR, 1986, Second Optional Protocol to the International Covenant on Civil and Political Rights, aiming at the abolition of the death penalty. https://www.ohchr.org/en/instruments-mechanisms/instruments/second-optional-protocol-international-covenant-civil-and.

European Union

Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases, OJ L 77, 27.3.1996.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, OJ L 157, 15.6.2016.

Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on copyright and related rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC, OJ L 130, 17.5.2019.

Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (recast), OJ L 172, 26.6.2019.

Directive (EU) 2022/2555 of the European Parliament and of the Council of 14 December 2022 on measures for a high common level of cybersecurity across the Union, amending Regulation (EU) No 910/2014 and Directive (EU) 2018/1972, and repealing Directive (EU) 2016/1148 (NIS 2 Directive), OJ L 333, 27.12.2022.

Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC, OJ L 117, 5.5.2017.

Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act), OJ L 151, 7.6.2019.

Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), OJ L 277, 27.10.2022.

Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), OJ L 265, 12.10.2022.

Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 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 (Artificial Intelligence Act), OJ L, 2024/1689, 12.7.2024.

Regulation (EU) 2024/2847 of the European Parliament and of the Council of 23 October 2024 on horizontal cybersecurity requirements for products with digital elements and amending Regulations (EU) No 168/2013 and (EU) 2019/1020 and Directive (EU) 2020/1828 (Cyber Resilience Act), OJ L, 2024/2847, 20.11.2024.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Regulation (EU) 2025/327 of the European Parliament and of the Council of 11 February 2025 on the European Health Data Space and amending Directive 2011/24/EU and Regulation (EU) 2024/2847, OJ L, 2025/327, 5.3.2025.

Case Law

CJEU C-524/06 Heinz Huber v. Germany, 16 December 2008, ECLI:EU:C:2008:724.

CJEU C-582/14, Breyer, 19 October 2016, ECLI:EU:C:2016:779.

CJEU T-557/20, SRB v EDPS, 16 April 2023, ECLI:EU:T:2023:219.

CJEU T-557/20, Single Resolution Board v. EDPS, 26 April 2023, ECLI:EU:T:2023:219.

CJEU C-131/12 Google Spain, 13 May 2024, ECLI:EU:C:2014:317.

ECtHR No. 42326/98, Odièvre v. France, 13 February 2003.

ECtHR No. 31871/96, Sommerfeld v. Germany, 8 July 2003.

ECtHR No. 77924/01, Albanese v. Italy, 23 March 2006.

ECtHR No. 17209/02, Zarb Adami v. Malta, 20 September 2006.

ECtHR No. 7508/02, L.L. v. Latvia, 10 October 2006.

ECtHR No. 73049/01, Anheuser-Busch Inc. v. Portugal, 11 January 2007.

ECtHR No. 30562/04, Marper v. UK, 4 December 2008.

ECtHR No. 13444/04, Glor v. Switzerland, 30 April 2009.

ECtHR No. 52019/07, L.H. v. Latvia, 29 April 2014.

ECtHR No. 6033/13, A.H. v. Russia, 17 January 2017.

Literature

Almeida, D. Shmarko, K. & Lomas, E., 2022, 'The ethics of facial recognition technologies, surveillance, and accountability in an age of artificial intelligence: a comparative analysis of US, EU, and UK regulatory frameworks', *AI and Ethics*, pp. 377-387.

Article 29 Data Protection Working Party, 2018, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, https://ec.europa.eu/newsroom/article29/items/612053.

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Article 29 Working Party, 2016, *Guidelines on Data Protection Officers ('DPOs') (WP 243)*, https://ec.europa.eu/newsroom/article29/items/612048.

Becker, R., e.a., 2022, 'Secondary Use of Personal Health Data: When Is It "Further Processing" Under the GDPR, and What Are the Implications for Data Controllers?', *European Journal of Health Law*.

BEUC, *Regulating AI To Protect The Consumer – Position Paper on the AI Act*, https://www.beuc.eu/sites/default/files/publications/beuc-x-2021-088_regulating_ai_to_protect_the_consumer.pdf.

Bickenbach, J., 2012*, Ethics, law and policy, Thousand Oaks*, pp. 19-24.

Bieker, F., Martin, N., Friedewald, M. & Hansen, M., 2018, 'Data Protection Impact Assessment: A Hands-On Tour of the GDPR's Most Practical Tool', *Privacy and Identity Management*, Springer, pp. 207-220.

Bitkina, O., Park, J. & Kim, H., 2023, 'Application of artificial intelligence in medical technologies: A systematic review of main trends', *Digital Health*.

Breen, S., e.a., 2020, 'GDPR: Is your consent valid?', *Business Information Review*, pp. 19-24.

Brittain, S., 2015, 'The Relationship Between the EU Charter of Fundamental Rights and the European Convention on Human Rights: an Originalist Analysis', *European Constitutional Law Review*, 11(3), pp. 482-511.

Buolamwini, J. & Gebru, T., 2018, 'Gender shades: Intersectional accuracy disparities in commercial gender classification', *Conference on Fairness, Accountability, and Transparency*, pp. 1-15.

Calvo, R.A., Peters, D., Vold, K. & Ryan, R.M., 2020, 'Supporting Human Autonomy in AI Systems: A Framework for Ethical Enquiry', *Ethics of Digital Well-Being. Philosophical Studies Series*, Springer, https://doi.org/10.1007/978-3-030-50585-1_2.

Cimina, V., 2021, 'The data protection concepts of 'controller', 'processor' and 'joint controllership' under Regulation (EU) 2018/1725', *ERA Forum*, pp. 639-654.

Costentini, A. e.a., 2025, *Assessing the Impact of Artificial Intelligence Systems on Fundamental Rights*, https://www.medialaws.eu/wp-content/uploads/2025/03/Assessing-the-Impact-of-Artificial-Intelligence-Systems-on-Fundamental-Rights.pdf.

Council of Europe, 2025, *A Convention to protect your rights and liberties*. https://www.coe.int/en/web/human-rights-convention/home.

Dastin, J., 2018, *Insight - Amazon scraps secret AI recruiting tool that showed bias against women*, https://www.reuters.com/article/world/insight-amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK0AG/.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Debes, R., 2023, 'Dignity', *Stanford Encyclopedia of Philosophy*, https://plato.stanford.edu/entries/dignity/.

Dewitte, P., 2024, 'AI Meets the GDPR, Navigating the Impact of Data Protection on AI Systems', *The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence*, pp. 133-157.

Douglas-Scott, S., 2015, 'The Relationship between the EU and the ECHR Five Years on from the Treaty of Lisbon', *Five Years Legally Binding Charter of Fundamental Rights*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2533207.

ECtHR, 2024, *Factsheet – Personal data protection*, https://www.echr.coe.int/documents/d/echr/FS_Data_ENG.

ECtHR, 2024, *Guide on Article 1 of Protocol No. 1 to the European Convention on Human Rights – Protection of property*. https://ks.echr.coe.int/documents/d/echr-ks/guide_art_1_protocol_1_eng.

EDPB, 2020, *Guidelines 05/2020 on consent under Regulation 2016/679*, https://www.edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_202005_consent_en.pdf.

EDPB, 2020, *Guidelines 07/2020 on the concepts of controller and processor in the GDPR, version 1.0*, https://edpb.europa.eu/sites/default/files/consultation/edpb_guidelines_202007_controllerprocessor_en.pdf.

EDPB, 2025, *Guidelines 01/2025 on Pseudonymisation*, https://www.edpb.europa.eu/system/files/2025-01/edpb_guidelines_202501_pseudonymisation_en.pdf.

EDPS, 2024, *Generative AI and the EUDPR. First EDPS Orientations for ensuring data protection compliance when using Generative AI systems*, https://www.edps.europa.eu/system/files/2024-06/24-06-03_genai_orientations_en.pdf.

EU, 2024, *European approach to artificial intelligence*, https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence.

EU, 2025, *Database protection*, https://europa.eu/youreurope/business/running-business/intellectual-property/database-protection/index_en.htm.

European Commission, 2010, *European Textbook on Ethics in Research*.

European Commission, 2018, *Ethics in Social Science and Humanities*, October 2018, https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020_ethics-soc-science-humanities_en.pdf, 15.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

European Commission, 2021, *Ethics and data protection*, https://ec.europa.eu/info/fundingtenders/opportunities/docs/2021-2027/horizon/guidance/ethics-and-data-protection_he_en.pdf.

European Commission, 2022, *Cyber Resilience Act*, https://digital-strategy.ec.europa.eu/en/library/cyber-resilience-act.

European Commission, 2022, *European Commission digital strategy Next generation digital Commission*, https://commission.europa.eu/document/download/70703206-2592-4175-b10d-12f97382094a_en?filename=C_2022_4388_1_EN_ACT.

European Commission, 2022, *Staff working document on data spaces*, https://digital-strategy.ec.europa.eu/en/library/staff-working-document-data-spaces.

European Commission, 2022, *The Digital Services Act*, https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/digital-services-act_en.

European Commission, 2024, *Data Act*, https://digital-strategy.ec.europa.eu/en/policies/data-act.

European Commission, 2024, *Data Governance Act explained*, https://digital-strategy.ec.europa.eu/en/policies/data-governance-act-explained.

European Commission, 2024, *New practical guide to the Data Governance Act*, https://digital-strategy.ec.europa.eu/en/library/new-practical-guide-data-governance-act.

European Commission, 2024, *The EU copyright legislation*, https://digital-strategy.ec.europa.eu/en/policies/copyright-legislation.

European Commission, 2025, *Common European Data Spaces*, https://digital-strategy.ec.europa.eu/en/policies/data-spaces.

European Commission, 2025, *European Health Data Space Regulation* (EHDS), https://health.ec.europa.eu/ehealth-digital-health-and-care/european-health-data-space-regulation-ehds_en.

European Commission, 2025, *Gatekeepers*, https://digital-markets-act.ec.europa.eu/gatekeepers_en.

European Commission, 2025, *Supervision of the designated very large online platforms and search engines under DSA*, https://digital-strategy.ec.europa.eu/en/policies/list-designated-vlops-and-vloses.

European Commission, 2025, *The Commission publishes guidelines on AI system definition to facilitate the first AI Act's rules application*, https://digital-

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

strategy.ec.europa.eu/en/library/commission-publishes-guidelines-ai-system-definition-facilitate-first-ai-acts-rules-application.

European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and The Committee of the Regions a European strategy for data*, 19 February 2020, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0066.

European Council, 2022, *Digital services package*, https://www.consilium.europa.eu/en/policies/digital-services-package/; European Commission, 2025, The Digital Services Act package, https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package.

European Council, 2024, *How the EU is strengthening its cybersecurity*, https://www.consilium.europa.eu/en/policies/cybersecurity/#act.

European Council, 2025, *Digital Services Act*, https://www.consilium.europa.eu/en/policies/digital-services-act/#act.

European IP Helpdesk, 2021, *Trade Secrets: Managing Confidential Business Information*, https://op.europa.eu/en/publication-detail/-/publication/5f1c6d8a-f015-11eb-a71c-01aa75ed71a1/language-en; EU, 2025, Database protection, https://europa.eu/youreurope/business/running-business/intellectual-property/database-protection/index_en.htm.

European Parliament, 2020, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, https://op.europa.eu/en/publication-detail/-/publication/dc544697-19b8-11ec-b4fe-01aa75ed71a1/language-en; FRA, 2018, Handbook on European data protection law, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG.

European Parliament, 2025, *Algorithmic discrimination under the AI Act and the GDPR*, https://www.europarl.europa.eu/RegData/etudes/ATAG/2025/769509/EPRS_ATA(2025)769509_EN.pdf.

Feldman, P., e.a., 2023, *Trapping LLM Hallucinations Using Tagged Context Prompts*, www.researchgate.net/publication/371490092_Trapping_LLM_Hallucinations_Using_Tagged_Context_Prompts/citations.

Finck M. & F. Pallas, F., 2020, 'They who must not be identified – distinguishing personal from non-personal data under the GDPR', *International Data Privacy Law* 2020, pp. 11-36.

Fink, M., *Human Oversight under Article 14 of the EU AI Act*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5147196.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Floridi L. & Taddeo, M., 2016, 'What is Data Ethics?', *Phil. Trans. R. Soc. A.*

FOD Economie Belgium, 2022, *European Directive on copyright and related rights in the Digital Single Market – transposition in Belgian law*, https://economie.fgov.be/en/themes/intellectual-property/intellectual-property-rights/copyright-and-related-rights/copyright/european-directive-copyright.

Foulsman, M., Hitchen, B. & Denley, A., 2019, *GDPR. How to Achieve and Maintain Compliance*, Routledge, pp. 21-23.

FRA, 2018, *#BigData: Discrimination in data-supported decision making*, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-focus-big-data_en.pdf.

FRA, 2018, *Handbook on European data protection law*, https://www.echr.coe.int/documents/d/echr/Handbook_data_protection_ENG.

FRA, 2019, *Data quality and artificial intelligence – mitigating bias and error to protect fundamental rights*, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2019-data-quality-and-ai_en.pdf.

FRA, 2020, *Getting the future right – Artificial intelligence and fundamental rights*, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-artificial-intelligence_en.pdf.

FRA, 2020, *Your rights matter: data protection and privacy*, https://fra.europa.eu/sites/default/files/fra_uploads/fra-2020-fundamental-rights-survey-data-protection-privacy_en.pdf.

FRA, 2025, *Article 1 – Human Dignity*, https://fra.europa.eu/en/eu-charter/article/1-human-dignity.

Gils, T., 2024, 'A Detailed Analysis of Article 50 of the EU's Artificial Intelligence Act', *The EU Artificial Intelligence (AI) Act: A Commentary*, Kluwer.

Global Alliance, 2024, GDPR *Brief: when are synthetic health data personal data?,* https://www.ga4gh.org/news_item/when-are-synthetic-health-data-personal-data/#:~:text=The%20key%20question%20is%20whether%20synthetic%20data%20fall,law%20as%20%E2%80%98personal%20data%E2%80%99%20%28Article%204%20%281%29%20GDPR%29.

GPAI, 2024, *Algorithmic Transparency in the Public Sector A state-of-the-art report of algorithmic transparency instruments*, https://wp.oecd.ai/app/uploads/2024/12/14-Algorithmic-Transparency-in-the-Public-Sector-A-state-of-the-art-report-of-algorithmic-transparency-instruments.pdf.

H.K., 2025, *The Limits of Human Oversight: What Alignment Research Reveals About the EU AI Act's Gaps*, https://www.linkedin.com/pulse/limits-human-oversight-what-alignment-research-eu-ai-acts-hernandez-zlclf/.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Hacker, P. & Passoth, J.H., 2020, 'Varieties of AI Explanations under the Law. From the GDPR to the AIA, and Beyond', *Lecture Notes on Artificial Intelligence 13200: beyond explainable AI*, Springer.

ISO, 2025, *AI Risk Assessments Under ISO/IEC 42001: A Practical Guide*, https://iso-docs.com/blogs/iso-42001-artificial-intelligence-management-system-aims/ai-risk-assessments-under-iso-iec-42001-a-practical-guide.

ISO, 2025, *Towards a trustworthy AI*, https://www.iso.org/news/ref2530.html#:~:text=To%20address%20issues%20of%20trust%20in%20artificial%20intelligence,concerns%20related%20to%20trustworthiness%20and%20provides%20practical%20solutions.

ISO/IEC, 2020, *ISO/IEC TR 24028:2020 Information technology — Artificial intelligence — Overview of trustworthiness in artificial intelligence*, https://www.iso.org/standard/77608.html.

ISO/IEC, 2023, *ISO/IEC 42001:2023 - AI management systems, https://www.iso.org/standard/81230.html*.

Jarovsky, L. *Can human really oversee AI?*, https://www.luizasnewsletter.com/p/can-humans-really-oversee-ai.

Karp, D.J., 2020, 'What is the responsibility to respect human rights? Reconsidering the 'respect, protect, and fulfill' framework', *International Treaty*. CUP, 12(1), pp. 83-108.

Keary, T., 2024, 'Hallucinations (Artificial Intelligence)', *Techopedia*, https://www.techopedia.com/definition/ai-hallucination#:~:text=Generative%20AI-driven%20chatbots%20can%20fabricate%20any%20factual%20information%2C,produce%20inaccurate%20information%2.

Kitting, M., e.a., 2024, 'Assessing trustworthy AI: Technical and legal perspectives of fairness in AI', *Computer Law & Security Review*.

Kloza, D., van Dijk, N. Casiraghi, S, Maymir, S.V., Roda, S., Tanas, A. & Konstantinou, I., 2020, 'Towards a method for data protection impact assessment: Making sense of GDPR requirements', https://doi.org/10.31228/osf.io%2Fes8bm.

Kordzadeh, N. & Ghasemaghaei, M., 2020, *Algorithmic bias: review, synthesis, and future research directions*, https://doi.org/10.1080/0960085X.2021.1927212.

KPMG, 2023, *Network & Information Security Directive (NIS2) NIS2 (EU) Directive Readiness - Levelling-up your IT and OT security capabilities in light of the NIS2*, https://assets.kpmg.com/content/dam/kpmg/pl/pdf/2023/10/kpmg-network-and-information-security-directive-nis2.pdf.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

KPMG, 2025, ISO/*IEC 42001: The latest AI management system standard*, https://kpmg.com/ch/en/insights/artificial-intelligence/iso-iec-42001.html.

Lambert, P., 2016, *The Data Protection Officer: Profession*, Rules, and Role, CRC Press, pp. 67-73.

Larson, J., Mattu, S., Kirchner, L. & Angwin, J., 2016, 'How We Analyzed the COMPAS Recidivism Algorithm', *ProPublica*.

Lievens E. & van der Hof, S., 2018, 'The importance of privacy by design and data protection impact assessments in strengthening protection of children's personal data under the GDPR', *Communications Law*, pp. 33-43.

Mennella, C., Maniscalco, U., De Pietro, G. & Esposito, M., 2024, 'Ethical and regulatory challenges of AI technologies in healthcare: A narrative review', *Heliyon*.

Naudts, L. & Vedder, A., 2024, 'Fairness and Artificial Intelligence', *The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence*, pp. 79-100.

NHB, 2022, 'Science must respect the dignity and rights of all humans'. *Natural Human Behaviour* 6, pp. 1029–1031.

NIS2 Compliant.org, 2024, *Comprehensive guide to the NIS 2 directive V.2.0*, https://nis2compliant.org/wp-content/uploads/2024/07/NIS-2-guide-1.pdf.

OHCHR, 2025, Background to the Covenant, https://www.ohchr.org/en/treaty-bodies/cescr/background-covenant; OHCHR, 2025, Committee on Economic, Social and Cultural Rights, https://www.ohchr.org/en/treaty-bodies/cescr.

OHCHR, 2025, *Human Rights Committee*. https://www.ohchr.org/en/treaty-bodies/ccpr.

OHCHR, *Civil and Political Rights: The Human Rights Committee, Fact Sheet No. 15*. https://www.ohchr.org/sites/default/files/Documents/Publications/FactSheet15rev.1en.pdf.

Onitiu, D., Wachter, S. & Mittelstadt, B., 'How AI Challenges the Medical Device Regulation: Patient Safety, Benefits, and Intended Uses'.

Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), OJ L 152, 3.6.2022.

Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act), OJ L, 2023/2854, 22.12.2023.

Renda, A., 2024, 'Europe: Toward a Policy Framework for Trustworthy AI', *The Oxford Handbook on Ethics of AI*. OUP.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

Riemann, R., 2025, *Synthetic Data*, https://www.edps.europa.eu/press-publications/publications/techsonar/synthetic-data_en.

Rochel, J., 2021, 'Ethics in the GDPR: A Blueprint for Applied Legal Theory', *International Data Privacy Law*, pp. 209-223.

Serio, M. e.a., 2023, 'Ethics in Legal Research', *Ethics in Research*. Springer, ALLEA, 2023.

Smith, M. & Miller, S., 2022, 'The ethical application of biometric facial recognition technology', *AI & Society*, pp. 167-175.

Smuha, N.A, 2019, 'The EU Approach to Ethics Guidelines for Trustworthy Artificial Intelligence', *Computer Law Review International*, pp. 97-106.

Smuha, N.A., e.a., 2021, *How the EU Can Achieve Legally Trustworthy AI: A Response to the European Commission's Proposal for an Artificial Intelligence Act,* https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3899991

Stadler, T., Oprisanu, B. & Troncoso, C., 'Synthetic Data -- Anonymisation Groundhog Day', arxiv.2011.07018.

Stalla-Bourdillon S. & Knight, A., 2017, 'Anonymous Data v. Personal Data — A False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data', *Wisconsin International Law Journal*.

Tasioulas, J., 2015, *Human Dignity and the Foundations of Human Rights*, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2557649#:~:text=This%20chapter%20investigates%20whether%20human%20rights%20are%20grounded,human%20dignity%3A%20the%20deontological%20and%20the%20personhood%20objections.

Tiedemann, P., 2023, 'Human Rights Concerning the Protection of Physical and Mental Integrity', *Philosophical Foundation of Human Rights*. Springer, pp. 141-158.

Truscott, J., Graham, A. & Powell, M.A., 2019, 'Ethical Considerations in Participatory Research with Young Children', *Participatory Research with Young Children. Educating the Young Child*, Springer, 2019.

UN, 2024, *AI has an environmental problem. Here's what the world can do about that*, https://www.unep.org/news-and-stories/story/ai-has-environmental-problem-heres-what-world-can-do-about.

UN, 2024, *Artificial intelligence (AI) end-to-end: The environmental impact of the full AI life cycle needs to be comprehensively assessed*, https://wedocs.unep.org/bitstream/handle/20.500.11822/46288/AI-Environmental-Impact-Issues-Note.pdf?sequence=3&isAllowed=y.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.

UNESCO, 2023, *The UNESCO Recommendation on The Ethics of AI: Shaping the Future of Our Societies*, https://www.unesco.nl/sites/default/files/inline-files/Unesco%20AI%20Brochure.pdf.

UNESCO, 2024, *Challenging systematic prejudices: an investigation into bias against women and girls in large language models*, https://unesdoc.unesco.org/ark:/48223/pf0000388971.

van Bekkum, M. & Borgesius F.Z., 2023, 'Using sensitive data to prevent discrimination by artificial intelligence: Does the GDPR need a new exception?', *Computer Law & Security Review*.

van Bekkum, M., 'Using sensitive data to de-bias AI systems: Article 10(5) of the EU AI act', *Computer Law & Security Review*.

van Kolfschooten, H. & van Oirschot, J., 'The EU Artificial Intelligence Act (2024): Implications for healthcare', *Health Policy*.

Van Noorden, R. 2020, 'The ethical questions that haunt facial-recognition research', *Nature*, https://www.nature.com/articles/d41586-020-03187-3.

Vanherpe, J., 2024, 'Artificial Intelligence and Intellectual Property Law', *The Cambridge Handbook of the Law, Ethics and Policy of Artificial Intelligence*, pp. 211-227.

Voigt P. & von dem Bussche, A., 2017, *The EU General Data Protection Regulation*, Springer, pp. 87-92.

Wach, K., e.a., 2023, 'The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT', *Entrepreneurial Business and Economics Review* 2023, http://dx.doi.org/10.15678/EBER.2023.110201.

Whittlestone, J., Nyrup, R., Alexandrova, A. & Cave, S., 2019, 'The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions.', *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, pp. 195–200.

Funded by the European Union

This project has received funding from the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101189689.