UNIVERSITY OF PIRAEUS

DEPARTMENT OF DIGITAL SYSTEMS

Postgraduate Program in

«LAW AND INFORMATION AND COMMUNICATION TECHNOLOGIES»

Academic year 2023-2024

LLP ERASMUS PROGRAMME: SAPIENZA UNIVERSITY OF ROME

MASTER THESIS

LOUKIA EMVALOMENOU (R.N.: MDI2315)

AI ACT AND PRIVACY

Supervisors:

S. Gritzalis (University of Piraeus), R. Acciai (Sapienza University of Rome)

Piraeus, 2025

Table of Contents

Contents

List of Abbreviations

AI	Artificial Intelligence
AIA	Artificial Intelligence Act
GDPR	General Data Protection Regulation
LED	Law Enforcement Directive
EU	European Union
ML	Machine Learning
DPAs	Data Protection Authorities
OECD	Organisation for Economic Co-operation and Development
EDPS	European Data Protection Supervisor
AIHLEG	High-Level Expert Group on Artificial Intelligence
ECHR	European Convention on Human Rights
CJEU	Court of Justice of the European Union
DPIA	Data Protection Impact Assessment
MSAs	Market Surveillance Authorities

Abstract

This thesis examines the risks that artificial intelligence (AI) poses to the fundamental right to privacy and provides a comprehensive analysis of the AI Act in light of these risks. It argues that the AI Act, while an important milestone for AI governance in the EU, does not sufficiently safeguard individuals' privacy. It concludes that it should be regarded as a complementary legislative piece to the General Data Protection Regulation (GDPR) and the Law Enforcement Directive (LED).

Key words: AI Act, personal data, artificial intelligence, AI, privacy, GDPR, AIA, LED, processing of personal data, personal data protection, AI training

Summary

The Artificial Intelligence Act (AI Act/Act/AIA), formally known as Regulation (EU) 2024/1689, is the EU's advanced legislative piece that establishes harmonized rules for the development, placing on the market and use of AI technologies within the EU, in effect as from 1 August 2024.

The Act came as a response to the alarming consequences of AI to fundamental rights and especially to the right to privacy. As will be further analysed, serious implications arise from the processing of huge amounts of personal data for the training of AI, in many cases without individuals' permission, but also from the use of AI for automated decision making. AI systems' decisions are often discriminatory against certain individuals or groups, reaching to conclusions without a clear logic behind them or producing fake results. To the worst end, AI can also be used for illicit purposes ranging from deepfakes dissemination to surveillance, biometric identification and social scoring based on users' profiling.

The Act clearly stipulates that it is without prejudice to existing EU data privacy legislation, i.e. the General Data Protection Regulation (GDPR), Regulation (EU) 2016/679 on the protection of natural persons with regard to the processing of personal data and Law Enforcement Directive 2016/680 (LED), focused on data privacy and security, which establishes specific principles to data processing, such as data minimization, purpose limitation, transparency, and grants respective rights to individuals. This thesis entails a short analysis of the AI Act in conjunction with the GDPR and LED.

The AIA follows a risk-based approach, defining four levels of risk posed by AI systems, with varying obligations for each one, namely: Unacceptable risk, high risk, limited risk and minimal or no risk. The Act establishes eight prohibited practices that pose unacceptable risks to the safety, livelihoods and rights of people, including harmful AI-based manipulation and deception, social scoring, emotion recognition in workplaces and education institutions, biometric categorisation and real-time remote biometric identification for law enforcement purposes in publicly accessible spaces. It also contains many obligations for high-risk AI systems, like security and risk assessment, which are often similar to the GDPR.

Nevertheless, the Act presents shortcomings when it comes to the protection of individual's privacy in the sphere of AI training and use. Indicatively, the scope of the prohibited practices is too narrow and it excludes many harmful practices from the general prohibition. This thesis aims to demonstrate that, even though the AIA is an important step to the regulation of AI in the EU, it fails to set adequate rules for the protection of the right to privacy. The GDPR is deemed the predominant and most adequate legislative piece for the protection of the right to privacy in this field. The Act should be seen as complementary to the GDPR, offering a structural approach and

governance, with a focus on product safety, while relying on the established data protection legislative pieces to safeguard individual rights.

Summary in Greek

α. Τα τελευταία χρόνια, η Τεχνητή Νοημοσύνη (ΤΝ) δεσπόζει στις τεχνολογικές, νομικές και πολιτικές εξελίξεις. Η ενσωμάτωσή της, τόσο στον δημόσιο όσο και στον ιδιωτικό τομέα, έχει οδηγήσει στην εκτεταμένη χρήση συστημάτων, όπως για παράδειγμα τα «chatbots», η «βιομετρική αναγνώριση» και τα «εργαλεία αυτοματοποιημένης λήψης αποφάσεων». Παρόλο που η χρήση της ΤΝ ενισχύει την αποδοτικότητα και συμβάλλει στην απλοποίηση καθημερινών λειτουργιών, εγείρει ταυτόχρονα σοβαρά ζητήματα όσον αφορά τα θεμελιώδη δικαιώματα, ιδίως το Δικαίωμα στην Ιδιωτικότητα και το Δικαίωμα στην Προστασία Προσωπικών Δεδομένων.

Τα συστήματα ΤΝ βασίζονται συχνά στην ευρείας κλίμακας επεξεργασία προσωπικών δεδομένων, δημιουργώντας κινδύνους επιτήρησης, κατάρτισης προφίλ και άλλων πρακτικών, που παραβιάζουν τα δικαιώματα των υποκειμένων. Ως απάντηση έναντι των κινδύνων αυτών, η Ευρωπαϊκή Ένωση θέσπισε τον Κανονισμό για την Τεχνητή Νοημοσύνη (Artificial Intelligence Act, AIA, Πράξη για την Τεχνητή Νοημοσύνης, ΠΤΝ), με σκοπό τη δημιουργία ενός εναρμονισμένου πλαισίου που ρυθμίζει την ΤΝ, διατηρώντας παράλληλα την καινοτομία. Η ΠΤΝ τέθηκε σε ισχύ τον Αύγουστο του 2024, ακολουθώντας μια προσέγγιση με βάση τον κίνδυνο που προκαλούν τα συστήματα ΤΝ (risk based approach). Στόχος της είναι η διασφάλιση της αξιοπιστίας και της διαφάνειας στην ανάπτυξη συστημάτων ΤΝ.

β. Η ΤΝ βασίζει τη λειτουργία της στην ανάλυση μεγάλων συνόλων δεδομένων προκειμένου να παράξει προβλέψεις, ταξινομήσεις και αυτοματοποιημένες αποφάσεις. Η παρούσα εργασία αναφέρεται στους κινδύνους που δημιουργεί η εξάρτηση των σύγχρονων συστημάτων ΤΝ, (ιδιαίτερα αυτών που χρησιμοποιούν μηχανική μάθηση (ML)), από τεράστιες ποσότητες προσωπικών δεδομένων με τα οποία τροφοδοτούνται για την εκπαίδευσή τους. Για παράδειγμα, η ανάπτυξη μεγάλων γλωσσικών μοντέλων (LLMs), περιλαμβάνει την αυτόματη συλλογή δεδομένων από το διαδίκτυο -συμπεριλαμβανομένων των κοινωνικών δικτύων και ιστολογίων- συχνά χωρίς τη γνώση ή συναίνεση των υποκειμένων. Αυτή η εξάρτηση από προσωπικά δεδομένα εγείρει νομικά ζητήματα που

άπτονται του Ενωσιακού Δικαίου περί ιδιωτικότητας και κυρίως του Γενικό Κανονισμό για την Προστασία των Δεδομένων (ΓΚΠΔ). Τα ζητήματα αυτά αφορούν ιδίως την παραβίαση των Αρχών της «ελαχιστοποίηση δεδομένων», της «νομιμότητας της επεξεργασίας» και της «διαφάνειας».

Η εκπαίδευση και η χρήση συστημάτων ΤΝ προκαλεί ουσιώδεις και πολύπλευρες συνέπειες που σχετίζονται με τα δικαιώματα του ατόμου στην Ιδιωτικότητα και στην Προστασία των Προσωπικών του Δεδομένων. Η εξάρτηση της ΤΝ από τη μαζική συλλογή και επεξεργασία προσωπικών δεδομένων εισάγει σύνθετους κινδύνους, όχι μόνο εξ αιτίας της εκτεταμένης συλλογής και επεξεργασίας δεδομένων αλλά και ως προς το είδος και την ευαίσθητη φύση των δεδομένων αυτών καθώς και ως προς την πιθανή χρήση των συστημάτων αυτών. Ένας από τους κινδύνους είναι η παράνομη συλλογή προσωπικών δεδομένων, συχνά μέσω διαδικτυακών πλατφορμών και κοινωνικών δικτύων, χωρίς συναίνεση, σαφή ενημέρωση ή εν γένει έγκυρη νομική βάση. Τα συστήματα ΤΝ που εκπαιδεύονται με τέτοιες πηγές είναι σε θέση να συμπεράνουν εξαιρετικά ευαίσθητα χαρακτηριστικά, όπως ο σεξουαλικός προσανατολισμός, οι θρησκευτικές πεποιθήσεις, η εθνοτική καταγωγή ή ακόμη και η συναισθηματική κατάσταση των υποκειμένων, δεδομένα τα οποία, σύμφωνα με τον ΓΚΠΔ, απαιτούν αυστηρή προστασία.

Ιδιαίτερα ανησυχητική είναι και η χρήση της ΤΝ για την κατάρτιση προφίλ και τη λήψη αυτοματοποιημένων αποφάσεων, καθώς μπορεί να οδηγήσει σε διακρίσεις ή να ενισχύσει κοινωνικές προκαταλήψεις. Τέτοια συστήματα χρησιμοποιούνται για την αξιολόγηση της πιστοληπτικής ικανότητας, την επιλογή υποψηφίων για εργασία ή την παροχή δημόσιων υπηρεσιών, συχνά χωρίς διαφάνεια ή δυνατότητα αμφισβήτησης. Σε πολλές περιπτώσεις, τα υποκείμενα των δεδομένων δεν γνωρίζουν καν, ότι μια απόφαση έχει ληφθεί από σύστημα ΤΝ, πόσο μάλλον ότι έχουν το δικαίωμα να την προσβάλουν. Αυτό θέτει σοβαρά ζητήματα δικαιοσύνης, λογοδοσίας και νομικής προστασίας.

Επιπλέον, η αδιαφάνεια πολλών συστημάτων ΤΝ, ειδικά αυτών που βασίζονται σε πολύπλοκα μοντέλα όπως τα νευρωνικά δίκτυα, επιτείνει το πρόβλημα. Σε πολλά συστήματα ΤΝ παρατηρείται το πρόβλημα του “μαύρου κουτιού”, δηλαδή της παραγωγής αποτελεσμάτων που δεν είναι ερμηνεύσιμα από ανθρώπους. Το γεγονός αυτό καθιστά τα

συστήματα αδιαφανή ως προς τον τρόπο λειτουργίας τους, οδηγώντας σε πρόκληση εσφαλμένων συμπερασμάτων ή την διαιώνιση διακρίσεων.

Οι κίνδυνοι για την ιδιωτικότητα εντείνονται περαιτέρω με την ανάπτυξη συστημάτων ΤΝ που πρόκειται να χρησιμοποιηθούν για βιομετρική ταυτοποίηση, κοινωνική βαθμολόγηση, προληπτική αστυνόμευση και αναγνώριση συναισθημάτων, τα οποία προσβάλλουν την ατομική αυτονομία. Η εφαρμογή τεχνολογιών συστημικής επιτήρησης εκδηλώνεται ήδη στην πράξη, ενώ παράλληλα, σε εμπορικό πεδίο, τα προσωπικά δεδομένα των χρηστών παράνομα μετατρέπονται σε προϊόντα και χρησιμοποιούνται για εξατομικευμένη στόχευση και εμπορική εκμετάλλευση.

Οι επιπτώσεις της εκπαίδευσης και χρήσης της ΤΝ στα ανθρώπινα δικαιώματα καταδεικνύουν με σαφήνεια ότι απαιτείται ισχυρό και εφαρμόσιμο ρυθμιστικό πλαίσιο για την ΤΝ.

Η ΠΤΝ υιοθετεί μια πολυεπίπεδη ρυθμιστική προσέγγιση με βάση τον κίνδυνο. Τα συστήματα ΤΝ ταξινομούνται ως «απαγορευμένα», «υψηλού κινδύνου» ή «ελάχιστου κινδύνου». Τα υψηλού κινδύνου συστήματα υπόκεινται σε αυστηρές απαιτήσεις συμμόρφωσης, όπως: θέσπιση υποχρεώσεων διαχείρισης κινδύνου, διαφάνειας, ανθρώπινης επίβλεψης καθώς και εποπτείας μετά τη διάθεση. Τα «γενικής χρήσης μοντέλα ΤΝ» υπόκεινται επίσης σε απαιτήσεις τεκμηρίωσης και συμμόρφωσης.

Η ΠΤΝ τονίζει την ανάγκη για "αξιόπιστη ΤΝ", ευθυγραμμίζεται με τις αρχές ανθρωποκεντρικής τεχνολογίας και αποτελεί ρυθμιστικό πλαίσιο για την ασφάλεια των προϊόντων, όμως δεν δημιουργεί δικαιώματα για τα υποκείμενα. Αν και η ΠΤΝ εισάγει νέα ρυθμιστικά επίπεδα για την ΤΝ, αναφέρεται ρητά στο κείμενό της ότι δεν θίγει την εφαρμογή των υφιστάμενων Κανονισμών Προστασίας της Ιδιωτικότητας , ιδιαίτερα του ΓΚΠΔ και της Οδηγίας για την Προστασία των Δεδομένων στο πλαίσιο επιβολής του νόμου (LED). Αυτό σημαίνει ότι η ΠΤΝ ρυθμίζει μεν τη λειτουργία και το σχεδιασμό των συστημάτων ΤΝ, αλλά δεν υποκαθιστά τις υποχρεώσεις για τη προστασία προσωπικών δεδομένων. Ωστόσο, η ΠΤΝ, με εξαίρεση το δικαίωμα επεξηγησιμότητας, δεν εισάγει άμεσα δικαιώματα αμφισβήτησης ή αναθεώρησης αποφάσεων που λαμβάνονται από συστήματα

ΤΝ, σε αντίθεση με τον ΓΚΠΔ. Η σχέση μεταξύ ΠΤΝ και ΓΚΠΔ είναι θεωρητικά συμπληρωματική, αλλά η απουσία σαφών μηχανισμών εναρμόνισης μπορεί να δημιουργήσει κενά στην προστασία και την εφαρμογή.

Το Άρθρο 5 της ΠΤΝ παραθέτει τις οκτώ απαγορευμένες πρακτικές ΤΝ: Χειριστικές ή παραπλανητικές τεχνικές, εκμετάλλευση τρωτών στοιχείων, αξιολόγηση κοινωνικής συμπεριφοράς, βάσεις δεδομένων αναγνώρισης προσώπου, πρόβλεψη τέλεσης αξιόποινης πράξης, συναγωγή συναισθημάτων, συστήματα βιομετρικής κατηγοριοποίησης, εξ αποστάσεως βιομετρική ταυτοποίηση. Όμως, το πεδίο εφαρμογής των απαγορευμένων πρακτικών είναι εν τέλει περιορισμένο καθώς η ΠΤΝ θέτει αυστηρές προϋποθέσεις ώστε να θεωρηθεί ένα σύστημα ΤΝ ως σύστημα «υψηλού κινδύνου» και άρα απαγορευμένο. Η παρούσα εργασία επικρίνει αυτήν την προσέγγιση ως υπερβολικά στενή, καθώς εξαιρεί πολλές δυνητικά επιβλαβείς τεχνολογίες.

δ. Η εργασία καταλήγει ότι οι διατάξεις για την προστασία της ιδιωτικότητας υπό την ΠΤΝ είναι δομικά ασθενέστερες σε σύγκριση με αυτές του ΓΚΠΔ και της LED. Αν και η ΠΤΝ απαιτεί διαφάνεια και ανθρώπινη επίβλεψη για τα υψηλού κινδύνου συστήματα, δεν θεσπίζει ατομικά δικαιώματα ανάλογα με εκείνα του ΓΚΠΔ. Η έλλειψη μηχανισμών προστασίας με βάση τα δικαιώματα σημαίνει ότι βασικές αρχές -όπως η Αρχή της «νομιμότητας», της «διαφάνειας» και της «λογοδοσίας»- δεν διασφαλίζονται επαρκώς στο πλαίσιο της ΠΤΝ. Επιπλέον, η τεχνική προσέγγιση της ΠΤΝ δεν ανταποκρίνεται πλήρως στις πολυδιάστατες απειλές της ΤΝ ως προς την ιδιωτικότητα. Συνεπώς, οι υφιστάμενοι νόμοι για την προστασία προσωπικών δεδομένων θεωρούνται πιο κατάλληλοι και αποτελεσματικοί για την προάσπιση των δικαιωμάτων των ατόμων στην ψηφιακή εποχή.

Συνοψίζοντας, η ΠΤΝ αποτελεί σημαντικό νομοθετικό βήμα για τη ρύθμιση της ΤΝ στην Ευρωπαϊκή Ένωση, όμως η συμβολή της στην προστασία της ιδιωτικότητας είναι περιορισμένη. Η έμφαση στη συστημική συμμόρφωση και την κατηγοριοποίηση βάσει κινδύνου δεν μπορεί να υποκαταστήσει την προστασία που προσφέρει το δικαιωματοκεντρικό πλαίσιο του ΓΚΠΔ και της LED. Η ΠΤΝ στερείται ισχυρών εγγυήσεων έναντι της συναγωγής συμπερασμάτων, της αυτοματοποιημένης λήψης αποφάσεων και της αδιαφανούς δημιουργίας προφίλ, ζητήματα που βρίσκονται στον πυρήνα των κινδύνων

ιδιωτικότητας που προκαλεί η ΤΝ. Συνεπώς, ο ΓΚΠΔ και η LED παραμένουν τα πιο ολοκληρωμένα και κατάλληλα νομικά εργαλεία για την προάσπιση της ιδιωτικότητας στην εποχή της τεχνητής νοημοσύνης.

1 INTRODUCTION

In recent years, the discussions around artificial intelligence (AI) have dominated the technological, legal, political and ethical landscapes. AI systems are being increasingly implemented by the public and private sector while automated decision making has become more prevalent and large language models (LLMs) have intruded our everyday lives. AI chatbots, biometric identification systems, virtual assistants, profiling and scoring systems are already integrated into our everyday lives.

Without doubt, AI simplifies many everyday processes enhancing our day-to-day activities in various ways and is making the completion of tasks faster, more efficient and accessible. Its integration into various fields continues to offer substantial benefits and enhance productivity.

However, alongside these advantages, it also introduces certain risks. One of the most pressing issues involves the risk it poses to individual's fundamental rights and freedoms. AI systems often process vast amounts of personal data, raising critical questions about privacy and data protection. As will be analyzed in chapter 3, the whole lifecycle of AI and its potential use threatens privacy in various ways.

In an attempt to regulate AI and protect against misuse without hindering innovation, the Regulation (EU) 2024/1689 of the European Parliament and the Council (Artificial Intelligence Act, AIA) entered into force on the 1st of August 2024, laying down a harmonized, comprehensive legal framework to regulate the use of AI across its Member States. Following a risk-based approach, the AIA prohibits certain AI systems that pose very serious risks to fundamental rights and sets rules for the placing on the market, the putting into service, and the use of AI systems and general purpose AI models and systems in the Union. Its aim is to set a "proportionate and effective set of binding rules for AI systems" (Recital 26, AIA) with great emphasis in the aspect of trustworthiness and ethical soundness of the AI systems.

The new Regulation has been in the epicenter of attention during the past years, since the publication of its Proposal. Many discussions have been going on around the AIA and its attempt to regulate this disruptive technology and address risks to fundamental rights like privacy and protection of personal data.

At the same time the AIA is without prejudice to existing EU data protection laws. (Article 2(7), AIA) The Regulation (EU) 2016/679 on the protection of personal data (known as the "General Data Protection Regulation") (GDPR) and the European Directive (EU) 2016/680 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties (known as the "Law Enforcement Directive") (LED) already regulate the collection and processing of personal data within the European Economic Area (EEA). While these Regulations don't specifically refer to AI, they take technological developments into account and provide rules regarding personal data processing and automated decision making.

This paper begins with an overview on the how AI works and its relationship with personal data. Then, the analysis turns to the dangers AI poses to privacy and its broader societal impact, referring to issues ranging from data collection for AI training, to risks associated with algorithmic bias, lack of transparency and misuse of AI. Relevant examples will be used to illustrate these concerns. Onwards, this thesis briefly analyses the AIA with the focus on the provisions related to privacy and the prohibited practices. It further examines the relevance and efficiency of the AIA to safeguard the fundamental right to privacy in relation to AI in comparison with the existing privacy Regulations, coming to the conclusion that the AIA does not sufficiently address the risks that AI pose to the fundamental right to privacy and data protection. Finally, this thesis argues that existing data protection laws, namely the GDPR and LED, are more adequate in addressing these issues.

2 AI SYSTEMS AND MODELS

2.1 The evolution of AI

"Any sufficiently advanced technology is indistinguishable from magic."

Arthur C. Clarke, Profiles of the future, 1962

The concept of AI is far from a new one. Since the early 19th century the public has been captivated by stories about human manufactured sentient beings and machines, like robots[1]. The term "artificial intelligence" was first introduced in 1955 by computer scientist John McCarthy at Dartmouth. (Wiggings and Jones, 2023) Despite the high hopes of success and the technological process in the following decades, the realization of AI remained elusive. In the mid-20th century, the era of the digital evolution, we experienced the emergence of the Internet, a rapid increase in computing power, a dramatic expansion in data storage capabilities, the growing influence of Big Data, and the rise of the Internet of Things. (Solove, 2024) Nevertheless, AI remained in the sphere of science fiction, until recently. In just a few years we have witnessed the transition from the so-called "AI winter", when the early expectations were not met, to spring, the era of growth and development. (Datalistnet, 2018). As the CNIL, the French Data Protection Authority, emphasizes "Algorithms and AI have come to represent new mythologies of our time". The rapid growth of AI is attributable to the enormous increase in computational power and access to huge amounts of data to train machine models. (Marengo, 2023)

2.2 AI Applications in Everyday Life

AI applications are now integrated into our daily lives shaping the way we live, work, and interact. AI is being used in a wide range of sectors, like healthcare, finance, retail, transportation, education and energy, marking significant breakthroughs. Some examples are the following;

[1] Frankenstein, The Modern Prometheus (1818), Mary Shelley, *I, Robot* (1950), Isaac Asimov, *2001: A Space Odyssey* (1968),

- Healthcare sector

Digital technologies and AI, particularly machine learning, are transforming medicine, medical research, and public health, with AI-based technologies now being used in healthcare services across various countries. According to the World Health Organization, the collection, analysis, and use of health data, including from clinical trials, laboratory results, and medical records, form the foundation of medical research and medical practice. More specifically, the most significant applications of AI in healthcare are in the diagnosis and prediction-based healthcare[2], in health research and drug development[3] and in public health and public health surveillance[4]. (WHO, 2021)

- Financial sector

Furthermore, a sector that is witnessing a profound transformation led by the ongoing technological revolution is the financial one[5], where enhanced automation leads to efficiency and productivity. By eliminating human errors and psychological biases, AI ensures accurate predictive analytics and trading strategies. It fosters business model innovation, particularly in

[2] "Currently, AI is being evaluated for use in radiological diagnosis in oncology (thoracic imaging, abdominal and pelvic imaging, colonoscopy, mammography, brain imaging and dose optimization for radiological treatment), in non-radiological applications (dermatology, pathology), in diagnosis of diabetic retinopathy, in ophthalmology and for RNA and DNA sequencing to guide immunotherapy. […] AI might be used to predict illness or major health events before they occur. For example, an AI technology could be adapted to assess the relative risk of disease, which could be used for prevention of lifestyle diseases. AI could assist in self-care, including through conversation agents (e.g. "chat bots"), health monitoring and risk prediction tools and technologies […] (WHO, 2021)

[3] "AI can nevertheless be applied to electronic health records for biomedical research, quality improvement and optimization of clinical care. […] as well as in drug discovery and throughout drug development to shorten the process and make it less expensive and more effective" (WHO, 2021)

[4] "AI has been used in public health surveillance for collecting evidence and using it to create mathematical models to make decisions and […] improve identification of disease outbreaks" (WHO, 2021)

[5] AI applications in finance include fraud detection, algorithmic and high-frequency trading, portfolio management, credit scoring, bankruptcy prediction, risk management, and regulatory compliance. Additionally, AI is used in stock market analysis, volatility forecasting, investor sentiment analysis, and managing credit risks and foreign exchange. (Bahoo, S. et al., 2024)

customized digital finance, which improves service efficiency and reduces costs. AI and machine learning enable rapid credit decisions, benefiting both lenders and consumers. (Bahoo, S. et al, 2024)

- Marketing and sales

Consumer interaction with AI has developed massively in the last decade. AI plays a crucial role in gathering valuable consumer insights about products and services, which is vital for both retaining existing customers and attracting new ones. The power of AI is evident in personalized service supply and predicted customer behavior analysis, which introduces a novel phase of marketing effectiveness[6] . (Labib, 2024)

The adoption of AI is expected to have significant implications for the subjects adopting them, and, more broadly, for the economy and society. (Bahoo, S. et al., 2024) It is anticipated that AI will contribute to global GDP growth, according to a 2017 study by PricewaterhouseCoopers (PwC) predicting an increase of up to 26% by 2030[7]. Additionally, companies that integrate AI technologies report improved performance. (Bahoo, S. et al, 2024) Nevertheless, these new practices pose significant risks for individual's fundamental rights.

[6] "The ability of AI to rapidly analyze extensive datasets empower businesses to extract invaluable insights, facilitating the development of targeted strategies that profoundly resonate with customers" (Labib, 2024) Marketers can use AI technology to identify emerging trends and forecast future patterns. With these insights, they can make informed decisions on budget allocation and determine the most effective target audiences. Thanks to the data collected and generated by its algorithms, marketers can also quickly determine what content to target customers and on which channel to employ at what moment. (Haleem et al, 2022)

[7] Predictions estimate a 26.1% GDP growth in China, 14.5% in North America, 9.9% in N. Europe, 11.5% in S. Europe and 10.4% in developed Asia.

2.3 Approaches to the definition of AI

2.3.1 AI definitions

While it seems that everyone is talking about AI, there seem to be many approaches as to the definition of it. AI can be defined as "a field of computer science that creates intelligent machines capable of performing cognitive tasks, such as reasoning, learning, taking action and speech recognition, which have been traditionally regarded as human tasks." (Investopedia, 2024) Nevertheless, as law professor Ryan Calo has stated: "There is no straightforward, consensus definition of artificial intelligence. AI is best understood as a set of techniques aimed at approximating some aspect of human or animal cognition using machines." This highlights the confusion, or at least the lack of precision surrounding the definition of AI. Given this ambiguity, it is essential, when considering the impact of AI on fundamental rights and freedoms, to develop a clear and comprehensive understanding of the concepts and functions at play. (Mitrou, 2018)

According to the Organisation for Economic Co-operation and Development (OECD) "An AI system is a machine-based system that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments. Different AI systems vary in their levels of autonomy and adaptiveness after deployment."

"Topics typically encompassed by the term "AI" and in the definition of an AI system include categories of techniques such as machine learning and knowledge-based approaches, and application areas such as computer vision, natural language processing, speech recognition, intelligent decision support systems, intelligent robotic systems, as well as the novel application of these tools to various domains." (OECD, 2024 A)

2.3.2 Machine Learning

Artificial intelligence, machine learning and deep learning are terms frequently used interchangeably even though they are conceptually imprecise. AI is an umbrella term that embraces many different types of machine learning. (Datalistnet, 2018) Machine learning can be

described as "a set of techniques and tools that allow computers to 'think' by creating mathematical algorithms based on accumulated data". (Datalistnet, 2018) With the ability to learn without being explicitly programmed, machine learning programs and techniques automatically enhance their performance through experience. This includes the design, analysis, development, and implementation of methods that enable a machine to function through a systematic process and accomplish complex tasks. It is important to note that the algorithm has the ability to define or adjust decision-making rules to manage new inputs. (Mitrou, 2018)

2.3.3 AI system definition according to the AI Act

The AIA gives a definition of an AI system. For the purposes of the AI Act "'AI system' means 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"[8] . This definition is intended to align with the OECD AI Principles. (Bird&Bird, 2024)

"A key characteristic of AI systems is their capability to infer. This capability to infer refers to the process of obtaining the outputs, such as predictions, content, recommendations, or decisions, which can influence physical and virtual environments, and to the capability of AI systems to derive models or algorithms, or both, from inputs or data. The techniques that enable inference while building an AI system include machine learning approaches that learn from data how to achieve certain objectives, and logic- and knowledge-based approaches that infer from encoded knowledge or symbolic representation of the task to be solved. The capacity of an AI system to infer transcends basic data processing by enabling learning, reasoning or modelling."[9] In contrast, traditional software that performs operations solely based on rules defined by humans, is not on its own an AI system. An AI system can be used independently or as a component of a

[8] Article 3 (1), AI Act.
[9] Recital 12, AI Act

product, regardless of whether the AI system is physically integrated into the product or serves the product's functionality without being integrated into it. (Bird&Bird, 2024)

2.4 AI and Data

The surge in artificial intelligence can largely be attributed to the rapid expansion of "datafication." (Mitrou, 2018). Technology enabled the digitalization of information, which converted analogue data into digital form. (Marengo, 2023) Over the first two decades of the 21st century, the business environment changed in many ways. The integration of machine learning with vast datasets has led to a significant increase in the development of products and services across both public and private sectors, spanning a wide variety of applications. (Mitrou, 2018)

As explained by Martin and Zimmermann, AI is built upon an ecosystem made up of three key technical components: data collection, data processing, and data outcome. The first step, data collection, involves information gathering, which can be either intentionally provided by individuals or acquired indirectly through digital footprints left by people during their routine activities (e.g., facial recognition technology in retail environments, internet browsing, fitness trackers). (Martin and Zimmermann, 2024) The collected data can be distinguished between structured and unstructured. Structured data refers to the data that are organized according to pre-defined models (eg. a relational database) and unstructured data when the organization is not known (eg piece of text or image).

The input can come in various formats. For a system designed solely for image recognition and analysis, the input data will naturally consist of images. For other tasks, the input may include text, speech, or numbers. Some systems use personal data[10], while others rely on data that cannot be traced back to individuals (Datalistnet, 2018) Algorithms improve and evolve with the input of large amounts of data. (Solove, 2024) Therefore, the effectiveness of AI algorithms depends on both the size and quality of the data used. (Rattan et al, 2022)

[10] Personal data means any information relating to an identified or identifiable natural person. (GDPR Article 4 (1))

The second element, data processing, is where the collected data are analyzed and interpreted with statistical and computational techniques. (Martin and Zimmermann, 2024) Machine learning techniques are designed to identify and mathematically represent patterns in data, (Rattan et al, 2022) making predictions or drawing inferences from them. (Solove, 2024)

The final element, data output, entails the delivery of processed information or an action. (Martin and Zimmermann, 2024) Algorithms are designed to anticipate outcomes and are used to take or support decisions vital to people's life with regard to finance, housing, employment, education and even justice and more. (Edwards and Veale, 2017)

2.5 AI and personal data

Although not all AI applications involve the processing of personal data, a large number of systems require extensive data about individuals or users, and as they operate, they increasingly extract more detailed information about them. (OECD, 2024 B) The data are collected from a variety of sources, including state or commercial databases, which may contain valuable information, such as equipment maintenance records, loan applications, financial transactions, and medical records. Machine learning systems then use these datasets to mine valuable insights, make predictions, or offer suggestions. (Solove, 2024) With the advancements of AI, all kinds of personal data can now be used to analyze, predict, and even shape human behavior. This capability transforms the data, along with the insights derived from its processing, into highly valuable assets that can be exploited for various purposes. (EPRS, 2020) As Giovanni Buttarelli states very accurately "Personal data and Artificial Intelligence are "a two -way street": personal data feeds AI and AI produces more inferred data".

While it is undeniable that these techniques bring new opportunities and mark impressive technological developments, the employment of personal data in the training of AI systems raises serious concerns. (OECD, 2024 B) The collection, processing and analysis of personal data required for using AI services or generated thereof, as will be further analyzed, may raise serious concerns about privacy and data protection (e.g. tracking and profiling), which are ultimately

inherent aspects of these services. (Solove, 2024) The use of AI to process vast amounts of personal data holds social importance; it can offer opportunities for enhanced societal understanding and improved governance, but it also carries the risk of contributing to the extremes of "surveillance capitalism" and the "surveillance state." (EPRS, 2020)

2.6 Generative AI and LLMs

Generative AI is a branch of artificial intelligence that specializes in creating new content, whether it be text, voice, images, or video, in response to prompts, based on the data the models have been trained on. Generative AI is based on machine learning (ML). (Solove, 2024) ML models use deep neural networks to mimic human intelligence by being exposed to data (training) and finding patterns that are then used to process previously unseen data. This allows the model to generalise based on probabilistic inference (i.e., informed guesses) rather than causal understanding. In order to achieve this successfully, deep neural networks need hundreds of thousands, millions, or even billions of examples, meaning that machine learning requires vast quantities of data. (OECD, 2024 A)

Since late 2022, generative AI, especially large language models (LLMs)[11] have captured global attention. At the center of this revolution is OpenAI's ChatGPT, which has become a key example of the potential and impact of these technologies. (Novelli et al, 2024) Known for their ability to generate coherent, natural-sounding sentences, LLMs can be used to create chatbot conversations, write articles, and even answer questions in a manner that feels more human-like compared to other AI technologies. (Blank, 2023) What sets LLMs apart is their ability to learn from patterns in existing text datasets, without needing labeled data or pre-programmed rules. This enables them to quickly process vast amounts of text data and respond to queries without the need for manual setup or programming. (Gokul, 2023)

[11] "Large language models (LLMs), are deep learning systems designed to process language. They are exposed to massive collections of texts, and their explicit training objective is to successfully predict the next word (or a missing word) in a sentence or paragraph (nowadays, additional objectives are included for further training). This objective is called 'language modeling'." (Blank, 2023)

Generative AI and LLMs raise important privacy issues due to extensive training on personal data, the memorization of training data, interactions with users and the output of AI products, (Novelli et al, 2024) receiving increasing regulatory attention as a result. (OECD, 2024 B)

2.7 Citations for chapter 2

Books

Wiggins, C. and Jones, M.L. (2023) *How data happened: A history from the age of reason to the age of algorithms.* W. W. Norton & Company.

Marengo, F. (2023) *Privacy and AI protecting individuals in the age of ai.* Independently published.

Articles

Solove, Daniel J. (2024) Artificial Intelligence and Privacy. 77 Florida Law Review (forthcoming Jan 2025), GWU Legal Studies Research Paper No. 2024-36, GWU Law School Public Law Research Paper No. 2024-36, Available at;

https://ssrn.com/abstract=4713111. or http://dx.doi.org/10.2139/ssrn.4713111.

Bahoo, S., Cucculelli, M., Goga, X. et al. (2024) Artificial intelligence in Finance: a comprehensive review through bibliometric and content analysis. *SN Bus Econ* 4, 23. https://doi.org/10.1007/s43546-023-00618-x.

Labib, E. (2024). Artificial intelligence in marketing: exploring current and future trends. Cogent Business & Management, 11(1). https://doi.org/10.1080/23311975.2024.2348728.

Haleem, A., Javaid M., Qadri M., Singh R., Suman R. (2022) Artificial intelligence (AI) applications for marketing: A literature-based study, International Journal of Intelligent Networks, Volume 3, Pages 119-132, ISSN 2666-6030, https://doi.org/10.1016/j.ijin.2022.08.005.

Calo, R. (2017) Artificial Intelligence Policy: A Primer and Roadmap. Available at SSRN: https://ssrn.com/abstract=3015350 or http://dx.doi.org/10.2139/ssrn.3015350.

Mitrou, L. (2018) Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'? Available at SSRN: https://ssrn.com/abstract=3386914 or http://dx.doi.org/10.2139/ssrn.3386914.

Martin K., Zimmermann J. (2024) Artificial intelligence and its implications for data privacy, Current Opinion in Psychology, Volume 58, 2024,

101829, ISSN 2352-250X, https://doi.org/10.1016/j.copsyc.2024.101829.

(https://www.sciencedirect.com/science/article/pii/S2352250X24000423)

Rattan P., Penrice D., Simonetto A. (2022) Artificial Intelligence and Machine Learning: What You Always Wanted to Know but Were Afraid to Ask, Gastro Hep Advances, Volume 1, Issue 1, Pages 70-78,ISSN 2772-5723, https://doi.org/10.1016/j.gastha.2021.11.001.

(https://www.sciencedirect.com/science/article/pii/S277257232100025X)

Edwards L. and Veale M. (2017) Slave to the algorithm? Why a 'right to an explanation' is probably not the remedy you are looking for. Duke Law and Technology Review, 16 (1). pp. 1-65, 19. https://scholarship.law.duke.edu/dltr/vol16/iss1/2/

Blank A. (2023) What are large language models supposed to model?,

Trends in Cognitive Sciences, Volume 27, Issue 11, Pages 987-989,ISSN 1364-6613, https://doi.org/10.1016/j.tics.2023.08.006.

(https://www.sciencedirect.com/science/article/pii/S1364661323002024)

Gokul A. (2023). LLMs and AI: Understanding Its Reach and Impact. Available at; https://www.researchgate.net/publication/370536412_LLMs_and_AI_Understanding_Its_Reach_and_Impact.

Novelli C., Casolari F., Hacker P., Spedicato G, Floridi L. (2024) Generative AI in EU law: Liability, privacy, intellectual property, and cybersecurity, Computer Law & Security Review, Volume 55,106066, ISSN 0267-3649, https://doi.org/10.1016/j.clsr.2024.106066. (https://www.sciencedirect.com/science/article/pii/S0267364924001328)

Publications/Reports by organizations

Datatilsynet, The Norwegian Data Protection Authority (January, 2018). *Artificial intelligence and privacy.* Oslo; The Norwegian Data Protection Authority.

CNIL, The French Data Protection Authority. (December, 2017). *How Can Humans Keep the Upper Hand? The Ethical Matters Raised by Algorithms and Artificial Intelligence*. Report on the Public Debate Led by the French Data Protection Authority (CNIL) as Part of the Ethical Discussion Assignment Set by the Digital Republic Bill. Available at; https://www.cnil.fr/sites/default/files/atoms/files/cnil_rapport_ai_gb_web.pdf.

World Health Organization (WHO). (2021) Ethics and governance of artificial intelligence for health: WHO guidance. Available at; https://hash.theacademy.co.ug/wp-content/uploads/2022/05/WHO-guidance-Ethics-and-Governance-of-AI-for-Health.pdf.

PricewaterhouseCoopers (PwC). (2017) Sizing the prize: PwC's Global AI Study—Exploiting the AI Revolution. Available at;

https://www.pwc.ch/en/publications/2017/pwc_global_ai_study_2017_en.pdf.

a. OECD (2024), "Explanatory memorandum on the updated OECD definition of an AI system", *OECD Artificial Intelligence Papers*, No. 8, OECD Publishing, Paris, https://doi.org/10.1787/623da898-en.
b. OECD (2024), "AI, data governance and privacy: Synergies and areas of international co-operation", *OECD Artificial Intelligence Papers*, No. 22, OECD Publishing, Paris, https://doi.org/10.1787/2476b1a4-en.

EPRS | European Parliamentary Research Service (2020) The impact of the General Data Protection Regulation (GDPR) on artificial intelligence. Brussels; European Parliamentary Research Service.

Websites

The Investopedia Team (April 09, 2024). Investopedia. Available at;

https://www.investopedia.com/terms/a/artificial-intelligence-ai.asp

[Accessed 20 Dec 2024].

Bird and Bird (2024). European Union Artificial Intelligence (EU AI Act) Guide. Available at;

https://www.twobirds.com/en/insights/2024/global/european-union-artificial-intelligence-act-guide [Accessed 20 Dec 2024].

Conferences

Buttarelli G., Privacy in an age of hyperconnectivity, Keynote speech to the Privacy and Security Conference 2016 Rust am Neusiedler See, 7 November 2016. "Personal data have increasingly become both the source and the target of AI applications", as expressed in the Council of Europe Consultative Committee Report on Artificial Intelligence and Data Protection, Strasbourg 17 September 2018.

3 AI AND PRIVACY IMPLICATIONS

3.1 Privacy and Data Protection

The right to make personal decisions, the right to keep one's personal information private, and the right to be left alone are essential components of the fundamental right to privacy. This right is widely recognized and safeguarded in numerous post-World War II human rights charters and is regarded as core principle of democracy[12]. Data privacy, according to the European Convention for the Protection of Human Rights (Article 8), "is primarily concerned with who has authorized access to collect, process, and potentially share one's personal data, and the extent to which one can exercise control over that access, including by opting out of data collection." The term has a broad scope, encompassing not only personal data but also any type of data that, if accessed by others, could be viewed as a violation of one's right to privacy and personal autonomy. (Stanford Institute for Human-Centered Artificial Intelligence (HAI), 2024)

Although closely related, privacy and data protection are generally regarded as two separate rights around the world. (EDPS, 2025) Data protection refers to the act of safeguarding individuals' personal information using a set of procedural rights, which includes ensuring that data is processed fairly, for specified purposes, and collected on the basis of one of the accepted bases for processing. (Article 8, CFR) (Stanford HAI, 2024) Nevertheless there are areas when the two overlap and complement each other. (Stanford HAI, 2024)

3.2 AI and Privacy concerns

Taking into consideration the nature of AI and the way AI algorithms function, by collecting data, processing them and producing outcomes based on them, the effects on personal privacy can be severe. (Martin and Zimmermann, 2024) The European Commission's White Paper (2020) listed 'loss of privacy, limitations to the right of freedom of expression, human dignity, discrimination

[12] Universal Declaration of Human Rights (Article 12), the European Convention of Human Rights (Article 8) and the European Charter of Fundamental Rights (CFR) (Article 7)

for instance in access to employment' as possible harms of AI. Different types of technologies are likely to have varying effects on informational privacy, decisional privacy, physical privacy and behavioral privacy. (Manheim and Kaplan, 2019) According to Manheim and Kaplan informational privacy is "the right to control the flow of our personal information", applicable to information we keep private and confidential information we share with others. Decisional privacy is "the right to make choices and decisions without intrusion or inspection". Behavioral privacy means "being able to do and act as one wants, free from unwanted observation or intrusion" and physical privacy "encompasses the rights to solitude, seclusion, and protection from unlawful searches and seizures". These notions of privacy have become a core feature of Western democracy as evidenced by their incorporation in fundamental documents and a vast body of statutory, common, and evidentiary laws. The most significant privacy risks posed by AI will be analyzed below.

3.3 Data Collection

As mentioned previously, AI technologies depend on increasing volumes of personal data to enhance the quality of their algorithms. As noted by Manheim and Kaplan (2019) "The more data collected the smarter, faster and more accurate the algorithms will be." Data can either be provided directly by citizens (volunteered data) or gathered by a third party (observed data). (Duberry, 2022, p. 105) Therefore, AI amplifies and significantly expands both the quantity and variety of data collected from consumers, while also motivating companies to breach the principles of data minimization and purpose limitation in their quest for more data. (Stanford HAI, 2024) Data collectors or third-party "cloud" storage providers store the extensive data gathered by IoT, surveillance, and tracking systems in various databases, which are then disclosed to government agencies and private entities. (Manheim and Kaplan, 2019)

Data for AI training are mainly collected in two ways: by data scraping from the internet or CCTV or by repurposing customer or user data. (Solove, 2024) It is considered that web scraping is the most efficient method for gathering large volumes of training data, particularly for generative AI applications. Text and media content available on publicly accessible websites serve as a vast

resource, providing natural language and image-text data essential for training generative AI models (Jayachandran and Arni, 2023). Many entities are engaging in data scraping practices from the internet, extracting data to feed their algorithms. (Solove, 2024)

This method compromises privacy, as highlighted by the class action against Clearview AI, a company known for developing one of the leading facial recognition tools, collecting billions of photos by scraping social media, online profiles, and photography websites. (Solove, 2024) The company allegedly used "stolen private information, including personally identifiable information, from hundreds of millions of internet users, including children of all ages, without their informed consent or knowledge." (CNN, 2023) In any case, "the fact that a person has published facial images of themselves on a social media platform does not mean that that person has given his or her consent for those images to be included in a facial recognition database" (European Commission, 2024) Regarding LLMs, reports suggest that LLM developers use vast amount of publicly accessible personal data to train their models. OpenAI follows the same practice to train ChatGPT. In fact, researchers were able to uncover parts of ChatGPT's training data, and discovered, among other things, phone and fax numbers, email and physical addresses, social media handles, names, birthdays, and even explicit content sourced from various areas of the internet. (Kuru, 2024)

Web scraping violates several widely recognized privacy principles outlined in laws, industry codes, and accepted standards, for example transparency regarding processing of personal data, clearly defining the objectives for processing personal data, providing information about third-party data sharing, seeking consent and granting individuals rights over their personal data. (Solove, 2024)

3.4 Legal basis for data collection

Even though companies have started to include in their privacy notices that they will use personal data for the development of AI, asking users to opt-in, the consent provided is often fictional. (Solove, 2024) There are many challenges with relying on consent because meaningful consent requires individuals to be fully informed about how their personal information will be used. If a

person does not fully understand how an AI system works or how their data will be used, their ability to provide informed consent is compromised. (Office of the Victorian Information Commissioner Office, 2021) Mitrou (2018) argues that it is highly questionable whether consent serves as an adequate legal basis due to the rapid and continuous evolution of technology and applications significantly impacting the foreseeability of future data uses based on the consent given.

Even if the data for training AI algorithms have been collected lawfully, their future use is unpredictable. In the initial stages of implementing an AI system, an organization may process personal data to train it, relying on a certain legal basis for this purpose. However, this legal basis could become irrelevant or inapplicable at a later stage, such as after the AI system is deployed or when the AI begins using personal data in new, unforeseen ways. The evolving nature of AI makes it difficult for organizations to predict how the system will use personal data in the future, often for purposes that are new, unexpected, or beyond the original intent of data collection. (Office of the Victorian Information Commissioner Office, 2021) (Mitrou, 2018)

3.5 Data Inference

Apart from the data directly collected, AI algorithms, regardless of whether they are trained with personal data or not, are designed to make inferences. Meaning that they are designed to recognize patterns, make predictions or draw conclusions from datasets that include information about numerous individuals, revealing information that individuals did not wish to share. (Martin and Zimmermann, 2024) (EDPB, 2024) AI models can draw conclusions and generate new information about both those individuals whose data have been used in the training process and other people as well. (European Data Protection Board, 2024) As Viljoen (2021) highlights, algorithms identify patterns among people, revealing significant insights into our biological, interpersonal, political, and economic relationships, posing notable privacy concerns. With inference and other methods of indirect collection, individuals may not always be aware that their personal information is being collected and therefore not have the opportunity to choose whether

or not to provide it (where there is a choice to do so or not). (Office of the Victorian Information Commissioner Office, 2021)

More precisely, they may be aware about the data gathered regarding them, but not the data generated about them. For example, the company Target had developed an algorithm that could identify pregnant women based on their shopping habits and then send them promoting material to inspire them to view Target as their preferred retailer for baby products. Target sent baby related advertisements to the father of a teenage girl who the algorithm had very correctly identified that she was pregnant. (Solove, 2024)

According to Solove (2024), privacy legislations often focus on the direct collection of data, rather than on the generation of data through inferential processes. While most privacy laws give individuals the right to rectify their data or consent to its collection, they generally do not provide sufficient means for individuals to challenge or correct inferences made from their data.

3.6 Anonymized Data

With AI's capabilities, not even data anonymization can ensure complete data protection. AI has the ability to re-identify anonymized data, data that have been stripped of personally identifiable information from collected data so the original source cannot be identified. By extracting relations from seemingly unrelated data, algorithms manage to make inferences. These include, for instance, AI's ability to generate comprehensive behavioral profiles from diverse datasets and to re-identify anonymized data, exposing our most intimate personal details to advertisers, governments, and strangers. (Manheim and Kaplan, 2019)

The European Data Protection Board (EDPB) (2024) considers that AI models trained on personal data cannot always be considered anonymous. Instead, the assessment of whether an AI model is anonymous should be made on a case-by-case basis, based on specific criteria. According to the EDPB, the AI model "may still retain the original information of those data, which may ultimately be extractable or otherwise obtained, directly or indirectly, from the model. Whenever information relating to identified or identifiable individuals whose personal data was used to

train the model may be obtained from an AI model with means reasonably likely to be used, it may be concluded that such a model is not anonymous". (EDPB, 2024)

3.7 Profiling and Automated Decision Making

Furthermore, AI algorithms are used not only to make inferences about past and present data but also to predict future events. (Solove, 2024) AI drives data analytics, allowing for predictive decision-making using consumers' financial, demographic, ethnic, racial, health, social, and various other data.

Machine learning and deep learning algorithms have the capability to categorize individuals based on specific parameters, leading to the formation of groups that share certain common characteristics. (Kosta, 2022) By merging, organizing, analyzing and correlating data from diverse databases, third parties create highly detailed profiles of persons that offer a wealth of valuable insights for anyone looking to influence or manipulate purchasing decisions and other choices (e.g., advertising targeting and election tampering) (Manheim and Kaplan, 2019)

Automated decision making, "instances in which algorithms or an AI are used to collect, process, and use data to make automated decisions" (Araujo et al, 2020) are increasingly being adopted in critical areas essential to the functioning of healthy democracies, such as public administration, law enforcement, education, healthcare, and hiring procedures. These systems are designed to enhance efficiency and boost the 'objectivity' of decision-making processes. (European Group on Ethics in Science and New Technologies, 2023)

In reality, these decisions are probabilities based on historical data, which frequently carry inherent biases, discrimination, inequalities, and privileges. If there is bias in the data, the algorithm will reflect it or possibly strengthen it, since the system uses feedback from these decisions to enhance its performance. As a result, automated decisions made using these algorithmic predictions often reinforce existing bias and disparities. (Araujo et al, 2020) (European Group on Ethics in Science and New Technologies, 2023) Especially considering AI's

ability to infer data, including sensitive personal data like sexual orientation, health conditions and race from public non-sensitive data, individual privacy is significantly compromised.

Some relevant examples where automated decisions are made are; assessing creditworthiness using loan repayment histories, forecasting the potential success of job applicants, making personalized news recommendations, determining criminal sentence lengths and overall limiting available choices. Algorithmic predictions pose challenges within current privacy law frameworks, with the primary mechanism for addressing them being individuals' right to correct such predictions. (Solove, 2024) Araujo et al. suggest that algorithmic predictions shape the future by predicting human behavior, thanks to the implementation of statistic methods. (2020)

3.8 Transparency

A key concern with ML is the "black box" phenomenon, the lack of transparency in how results are produced, as models often generate outcomes without explanation. This raises the question of whether it's possible to analyze the model to understand how it reached a specific result. (Datatilsynet, 2018) Transparency can be viewed from a scientific and legal perspective. According to Almada, scientific transparency usually refers to "the visibility of the mechanisms that produce the computational results produced by an AI system", essential to ensure the scientific validity of the results generated by an AI system, which is often a requirement for its lawful use. On the other hand, legal transparency is "a form of disclosure" referring to what, when and to whom is disclosed. (Almada, 2023)

Although various laws seek to regulate automated decisions through transparency requirements, such transparency often lacks true clarity. Due to the complexity of AI algorithms, simply knowing the logic behind these decisions is inadequate, as they depend on personal data from millions, which cannot be shared without violating privacy regulations. Without access to the data used to train these algorithms, evaluating certain automated decisions becomes a difficult, if not impossible, task. (Solove, 2024) This creates significant challenges regarding the legitimacy and accountability of decisions made using AI. (Duberry, 2022, p. 22) In fact, many of the largest

companies developing generative AI systems have been unresponsive to public inquiries regarding the sources of their data and the procedures they follow to remove personally identifiable information and other sensitive elements from their training data. (Stanford HAI, 2024)

3.9 Biometric identification

By processing biometric data, AI has contributed significantly to the advancement of biometrics and more specifically biometric identification. Biometric data are "personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, such as facial images or dactyloscopic data". (Article 3(34), AIA) DNA, retinal pattern, iris pattern, voice matching, body parts shape and other behavioral characteristics like gait and typing (Thomson Reuters, n.d.) are used to establish "the identity of a natural person by comparing biometric data of that individual to biometric data of individuals stored in a database". (Article 3(35), AIA) As technology evolves, the range of biometric data that can be obtained from individuals will also expand. (Holistic AI, n.d.)

Biometric identification offers the capability of real-time remote or post-remote identification in publicly accessible spaces, a tactic used particularly for law enforcement purposes and private security. Post remote identification could be utilized to identify individuals in a feed from public spaces hours, weeks, or even months after it was captured. It is used for example in various countries (e.g. US, Austria) in the context of a political protest, to identify the participants. (Euronews, 2023) Such systems are able to identify thousands of individuals in only a few hours, "having a severe effect on the populations' (reasonable) expectation of being anonymous in public spaces, resulting in a direct negative effect on the exercise of freedom of expression of assembly of association as well as freedom of movement". (EDPB and EDPS, 2021). The highly intrusive nature of such systems in combination with the lack of transparency they present renders them a threat to privacy.

3.10 Emotion recognition

Apart from inferencing information from data performing biometric identification, ML systems ares capable of recognizing emotions by inferring "an individual's inner affective state based on traits such as facial muscle movements, vocal tone, body movements, and other biometric signals. They use machine learning to analyse facial expressions and other biometric data and subsequently infer a person's emotional state". This process requires the collection of massive amounts of sensitive personal data. (ARTICLE 19, 2021) Such systems can potentially be used in education (e.g. to monitor students' attention), employment (e.g. identify uninterested candidates in a job interview), healthcare (e.g. detect autism and predict disorders), marketing (e.g. analyze customer's emotions while shopping), crime detection (e.g. lie detectors and smart boarders control) etc. (European Data Protection Supervisor, 2021)

According to the EDPS, the results may not always be accurate since facial expressions and expressions of emotion vary amongst people and cultures. In any case, since the results of facial emotion recognition, regardless of accuracy limitations, are usually treated as facts and are used as input to processes affecting a data subject's life, instead of triggering an evaluation to discover more about their situation in the specific context, are threatening human autonomy and privacy. (2021)

3.11 Deepfakes, fake news and hallucinations

AI has enabled the creation and distribution of Deepfakes, a generative AI technique that employs deep learning to create realistic yet artificial images, audio, or videos. Even though they have legitimate and transformative applications in fields such as marketing, historical reconstruction, education, and training, they present, however, significant risks, (e.g., generation of non-consensual deepfake pornography). The growing accessibility of generative AI tools has lowered the technical barriers of creating such content, broadening their potential use by malicious actors. Without clear labeling practices or advanced content provenance tools, it becomes challenging to determine when media has been generated by AI. (Association of Southeast Asian Nations, 2025)

Malicious use of deepfakes and fake content may include impersonation and dissemination of misinformation and disinformation causing harm to someone's reputation, or the very real potential of manipulating public opinion or even destabilizing democratic processes by spreading biased or false information targeted at the most vulnerable or impressionable audiences. (Uuk et al., 2024) Fake news can easily manipulate public opinion and have a severe impact on decision making. (Homo Digitalis, 2024)

In addition, LLMs like ChatGPT often give false information about people without providing a way to correct them. These so called "AI hallucinations" can vary from innocent mistakes to defamatory lies and can seriously damage a persons reputation. "There are multiple media reports about made up sexual harassment scandals, false bribery accusations and alleged child molestation – which already resulted in lawsuits against OpenAI". (noyb, 2025)

3.12 Surveillance Capitalism

Considering that digital technologies are nowadays utilized for various tasks, for example communicating through online platforms, making online purchases, searching for information, applying for a job, checking our bank account, (Duberry, 2022, p. 93) user data are continuously collected. On all browsing devices, small files known as "cookies" monitor every online activity from their position on our hard drive, subsequently transmitting this data to remote servers over the internet. (Duberry, 2022, p. 101) AI and machine learning (ML) enable both private and public entities to analyze and interpret the vast amounts of data collected, known as "big data", extracting valuable information. (Duberry, 2022, p. 94)

Particularly in the case of online platforms, a small number of tech companies known as GAMAM (Google, Amazon, Meta, Apple, and Microsoft) concentrate the majority of data collection, analysis, and monetization capabilities, and therefore, the majority of the benefits derived from this data. Nevertheless, surveillance on social media platforms is conducted not only by the platforms themselves and their ML algorithms (MLAs), but also by various other private entities, such as digital marketing companies, brands, data brokers, political parties, and numerous other users. (Duberry, 2022, p. 97)

Surveillance is ingrained within social media platforms whose MLAs are designed to keep users engaged and online so that the platforms and others can fulfill their economic interests by influencing and directing users, collecting more data and using their attention time to display personalized ads. (Duberry, 2022, p. 97) The internet, along with the various applications it supports, including social media platforms, has enabled digital marketers (including political) to track individual behavior in real time. (Duberry, 2022, p. 90) Manipulation and deception by platforms is one of the most dangerous aspects of AI. (Homo Digitalis, 2024) Behavioral advertising interferes with decisional privacy as well as informational privacy allowing for consumer and voter manipulation (Manheim and Kaplan, 2019) (Stanford HAI, 2024)

The Cambridge Analytica scandal that came to light in 2018, consists of a major data privacy breach from Cambridge Analytica, a consulting firm managing the political advertising of US elections candidate Donald Trump. The company illegally harvested massive amounts of Facebook user data and used AI to distribute targeted advertising in favor of candidate Trump. It managed to persuade voters by showing them advertisements that were tailored exactly to what would resonate with their personalities. (Politico, 2020)

As Zouboff contends "'big data' is above all the foundational component in a deeply intentional and highly consequential new logic of accumulation that I call surveillance capitalism. This new form of information capitalism aims to predict and modify human behavior as a means to produce revenue and market control". (2015, p. 75) "The model's fundamental characteristics are: aggregating vast amounts of data on people, using it to infer incredibly detailed profiles on their lives and behavior, and monetizing it by selling these predictions to others such as advertisers." Cambridge Analytica simply deployed the same basic model to target voters rather than consumers. (Amnesty International, 2019)

3.13 Surveillance State

As analyzed above, AI technologies have made big data processing, profiling, monitoring and tracking easier, cheaper and more accurate. (Duberry, 2022, p. 93) These sophisticated

identification techniques are being implemented by both the public and private sector, consequently enabling mass surveillance with serious implications for privacy and democracy.

Surveillance techniques powered by AI, ML and big data are increasingly adopted by governments as well, since they provide enhanced capabilities for state surveillance, enabling both law enforcement authorities and security and intelligence agencies to monitor and track activities more effectively. The deployment of AI in surveillance therefore marks the beginning of a new era of state monitoring, specifically referred to as algorithmic (mass/bulk) surveillance. (Kosta, 2022). Government agencies are increasingly using advanced methods to monitor communications, access electronic devices through hacking and surveil public spaces online and offline, with the help of biometric and artificial intelligence-based technologies. (United Nations Human Rights Council, 2024)

With the stance of protecting public security, states are implementing AI powered surveillance techniques used to prevent crime, police borders, monitor public behavior, and identify suspected terrorists. Some examples may be license plate readers, drones, smart policing tools (Francis, 2024) (Manheim and Kaplan, 2019), biometric identification and emotion recognition systems which can identify an individual and track their actions, social interactions, and potentially every word and behavior, in real time. (Solove, 2024) (Duberry, 2022, p. 44) The capabilities of AI powered surveillance cameras have increased dramatically, and they continue to grow, enabling precise identification. In fact, by 2018, 40 out of the top 50 countries with the highest military spending had implemented AI surveillance technologies. (Francis, 2024) According to a UN report, surveillance cameras in use globally exceeded one billion in 2021. (United Nations Human Rights Council, 2022)

These developments typically take place in the context of emerging identity systems and growing biometric databases. In many countries, identity systems are connected to large central repositories of personal data, including biometric details like fingerprints, facial features, iris scans, and DNA. Additionally, these databases are often interconnected and accessible for searching by various agencies. As a result, tracking individuals' locations has become increasingly easier. (United Nations Human Rights Council, 2022)

Law enforcement agencies and governmental organizations from 25 countries have used Clearview AI's facial recognition tool for biometric mass surveillance. (BuzzFeed News, 2021)

According to Duberry, states are collecting data which then analyze using ML algorithms, either created internally by their intelligence agencies or supplied by private military firms to identify individuals or to group them. (2022, p. 100) These algorithms, after profiling and grouping individuals, can detect violations reactively or pre-emptively, meaning, identify a violation in real time or predict a future unlawful behavior by inference. (Kosta, 2022) This person-based predictive policing (crime forecasting, crime prediction) is a practice already used in the US but also European countries like the UK, Germany, Netherlands, Estonia and Romania. The biggest issue regarding these systems is the bias in the result that appears due to biased training data sets. As a result, individuals from minority groups are more likely to be perceived as potential criminals, often without a robust basis. (Homo Digitalis, 2024) Kosta argues that this type of systemic surveillance subjects individuals to state surveillance without any real suspicion. Since they are unaware of these practices, they cannot comprehend how decisions about them are being made and object to them. (2022)

Systematic surveillance of individuals in both public and online spaces, especially when combined with methods to analyze and link the gathered data to other sources, represents a serious infringement on the right to privacy. This can also have severe negative impacts on the enjoyment of other fundamental human rights, especially for minorities and marginalized communities which are disproportionately targeted. (United Nations Human Rights Council, 2022) Extensive algorithmic profiling has altered societal norms and has established the expectation that people's data will be collected at every turn. The feeling of powerlessness in doing anything about it, and the lack of transparency regarding how people's data is used or decisions are made about them all contribute to a growing sense of inevitability that our privacy has already been compromised. (Stanford HAI, 2024)

3.14 Social scoring

In addition, AI can be used to evaluate individuals based on some social characteristics. These "social rating" or "social scoring" AI systems assess or categorize individuals or groups based on a range of data points, known or inferred, related to their social behavior in different contexts, or anticipated personal traits and characteristics over defined periods (AI Act, 2024) This score may affect individual's access to services, occupation or other opportunities. (Homo Digitalis, 2024)

In some societies social scoring is approved as it aligns with their social principles and values. In China, for example, a social credit system (SCS) is being implemented to assess the reliability of its citizens. The Chinese government is using various methods and policies to surveil citizens online and offline by collecting personal information like ID numbers, employment records, and education data but also more sophisticated methods like big data analytics, internet censorship and CCTV. In fact, China has installed more than 20 million street cameras which enable the government to not only capture criminal activity but also to identify and track individuals through facial recognition. (Liang et al., 2018) Based on this profile citizens are granted privileges or are excluded from benefits (e.g. booking a plane ticket to leave the country, having insurance, renting a house). (Kostopoulou, 2023)

Nevertheless, there are social scoring systems operating in Europe as well. The SyrI system, for example, is an algorithm used in the Netherlands, designed to identify potential social welfare fraud. However, the algorithm allegedly did a cross reference of citizen data from various databases and sources processing a lot more data than the necessary and evaluated citizens based on unrelated data. (Homo Digitalis, 2024)The Hague's District Court in its 2020 judgment held that "the use of SyRI entailed an interference with the right to respect for private life, which was unnecessary and disproportionate to the purpose of combating fraud" and highlighted the "absence of transparency and information about how the AI system worked" and the "difficulties of substantiating the discriminatory effects of the use of algorithmic systems". (Rachovitsa and Johann, 2022)

3.15 The importance of privacy

In conclusion, it is made clear that the impact of AI to the fundamental right to privacy is severe and concerning. The arguments "I have nothing to hide" or "only criminals should be concerned with their privacy" are often used to justify surveillance, collection of personal data and privacy violations in general. This is far from accurate. (Cofone, 2020) Apart from the repercussions from the use of AI highlighted above, like data exploitation by companies, consumer manipulation for profit maximization and discrimination, surveillance creates a "chilling effect", the notion that when one is being watched, their behavior inevitably changes. "The gaze of the 'watcher' is internalized by us, the people, and comes to shape what we do, how we think and ultimately who we are. Surveillance curtails our autonomy." (European Group on Ethics in Science and New Technologies, 2023). AI-powered surveillance practices can deeply influence how individuals think and behave, as well as other personal rights, such as freedom of expression or association. This is crucial for the active involvement of civil society in policy-making. (Duberry, 2022, p. 95)

Therefore, it is now to be examined whether the long awaited AI Act takes into consideration the serious implications of AI to the fundamental right to privacy and regulates it effectively.

3.16 Citations for chapter 3

Books

Duberry, J. (2022) *Artificial intelligence and democracy: Risks and Promises of AI-Mediated Citizen-Government Relations*. Cheltenham, UK, Northampton, USA: Edward Elgar Publishing.

Articles

Manheim, K. and Kaplan, L. (2019) 'Artificial Intelligence: Risks to Privacy and Democracy', *Yale Journal of Law & Technology*, Volume 21, (106) pp 116-131. Available at; https://yjolt.org/artificial-intelligence-risks-privacy-and-democracy [Accessed 7 February 2025].

Martin K., Zimmermann J. (2024) 'Artificial intelligence and its implications for data privacy', *Current Opinion in Psychology*, Volume 58, 101829, pp 1-3, ISSN 2352-250X, https://doi.org/10.1016/j.copsyc.2024.101829.

(https://www.sciencedirect.com/science/article/pii/S2352250X24000423)

Solove, D. J. (2024) 'Artificial Intelligence and Privacy'. *Florida Law Review* (forthcoming Jan 2025), GWU Legal Studies Research Paper No. 2024-36, GWU Law School Public Law Research Paper No. 2024-36, Available at;

https://ssrn.com/abstract=4713111. or http://dx.doi.org/10.2139/ssrn.4713111 [Accessed 7 February 2025].

Viljoen, S., (2021) 'A Relational Theory for Data Governance', *The Yale Law Journal*, Volume 131, (2) pp 370-381. Available at; https://www.yalelawjournal.org/feature/a-relational-theory-of-data-governance [Accessed 7 February 2025].

Araujo, T., Helberger, N., Kruikemeier, S., & de Vreese, C. H. (2020). 'In AI we trust? Perceptions about automated decision-making by artificial intelligence', *AI & Society*, 35(3), 611-623. Available at; https://doi.org/10.1007/s00146-019-00931-w [Accessed 7 February 2025].

Mitrou, L. (2018) 'Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?' University of the Aegean, Unpublished. Available at; SRN: https://ssrn.com/abstract=3386914 or http://dx.doi.org/10.2139/ssrn.3386914 [Accessed 11 February 2025].

Almada, M. (2023) 'Governing the Black Box of Artificial Intelligence', Universite du Luxembourg, Unpublished. Available at; SSRN: https://ssrn.com/abstract=4587609 or http://dx.doi.org/10.2139/ssrn.4587609 [Accessed 7 February 2025].

Jayachandran, J. and Arni, V. (2023) Traversing the Ethical Landscape of Data Scraping for AI, University of Maryland, Unpublished Available at SSRN: https://ssrn.com/abstract=4666354 or http://dx.doi.org/10.2139/ssrn.4666354.

Kuru T., (2024) 'Lawfulness of the mass processing of publicly accessible online data to train large language models', *International Data Privacy Law,* Vol. 00, No. 0, pp 1-25. Available at; https://doi.org/10.1093/idpl/ipae013 [Accessed 20 February 2024].

Cofone, I. (2020) 'Nothing to Hide, but Something to Lose'. *University of Toronto Law Journal*, Vol 70 pp 3-5, 43 Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3327646 [Accessed 21 February 2025].

Zuboff, S. (2015) 'Big other: Surveillance capitalism and the prospects of an information civilization'. *Journal of Information Technology*, 30(1), p 75.

Kosta E., (2022) 'Algorithmic state surveillance: Challenging the notion of agency in human rights'. *Regulation and Governance, Vol. 16, No 1, pp 212-224.* Available at: https://doi.org/10.1111/rego.12331 [Accessed 21 February 2025].

Uuk R. et al. (2024) "A Taxonomy of Systemic Risks from General-Purpose AI", Unpublished. Available at: SSRN: https://ssrn.com/abstract=5030173 or http://dx.doi.org/10.2139/ssrn.5030173 [Accessed 21 February 2025].

Liang, F., Das, V., Kostyuk, N. and Hussain, M.M. (2018), 'Constructing a Data-Driven Society: China's Social Credit System as a State Surveillance Infrastructure', *Policy & Internet*, 10(3), pp. 263-288. Available at: https://www.researchgate.net/publication/326817957 [Accessed 20 March 2025].

Rachovitsa, A. and Johann, N. (2022) 'The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case', *Human Rights Law Review*, 22(2), ngac010. Available at: https://doi.org/10.1093/hrlr/ngac010 [Accessed 27 March 2025].

White paper

Stanford University Institute for Human-Centered Artificial Intelligence (HAI), (2024) Rethinking Privacy in the AI Era; Policy Provocations for a Data-Centric World.

Available at; https://hai.stanford.edu/white-paper-rethinking-privacy-ai-era-policy-provocations-data-centric-world [Accessed 10 February 2025].

Webpages

European Data Protection Supervisor (2025). ***Data Protection***. Available at; https://www.edps.europa.eu/data-protection/data-protection_en [Accessed 7 February 2025].

CNN (2023). ***OpenAI, maker of ChatGPT, hit with proposed class action lawsuit alleging it stole people's data.*** Available at; https://edition.cnn.com/2023/06/28/tech/openai-chatgpt-microsoft-data-sued/index.html [Accessed 1 February 2025].

Amnesty International (2019). ***'The Great Hack': Cambridge Analytica is just the tip of the iceberg.*** Available at; https://www.amnesty.org/en/latest/news/2019/07/the-great-hack-facebook-cambridge-analytica/ [Accessed 22 March 2025].

Politico (2020). ***AI Decoded: How Cambridge Analytica Used AI, No, Google Didn't Call for a Ban on Face Recognition, Restricting AI Exports***. Available at; https://www.politico.eu/newsletter/ai-decoded/politico-ai-decoded-how-cambridge-analytica-used-ai-no-google-didnt-call-for-a-ban-on-face-recognition-restricting-ai-exports/ [Accessed 25 March 2025].

Thomson Reuters (n.d.). ***The Basics, Usage, and Privacy Concerns of Biometric Data.*** Available at: https://legal.thomsonreuters.com/en/insights/articles/the-basics-usage-and-privacy-concerns-of-biometric-data [Accessed 25 March 2025].

European Data Protection Supervisor (2021). ***Facial Emotion Recognition.*** Available at: https://www.edps.europa.eu/data-protection/our-work/publications/techdispatch/techdispatch-12021-facial-emotion-recognition_en [Accessed 26 March 2025].

Holistic AI (n.d.). ***Regulation of AI in Biometrics***. Available at: https://www.holisticai.com/papers/regulation-of-ai-in-biometrics [Accessed 26 March 2025].

BuzzFeed News (2021). ***Clearview AI's International Search Table***. Available at: https://www.buzzfeednews.com/article/ryanmac/clearview-ai-international-search-table [Accessed 26 March 2025].

Euronews (2023). ***Retrospective Facial Recognition Surveillance Conceals Human Rights Abuses in Plain Sight.*** Available at: https://www.euronews.com/2023/04/14/retrospective-facial-recognition-surveillance-conceals-human-rights-abuses-in-plain-sight [Accessed 7 April 2025].

noyb (2025). ***AI Hallucinations: ChatGPT Created Fake Child Murderer.*** Available at: https://noyb.eu/en/ai-hallucinations-chatgpt-created-fake-child-murderer [Accessed 7 April 2025].

Publications

Office of the Victorian Information Commissioner Office (2021) *Artificial Intelligence; Understanding Privacy Obligations.* Melbourne; Office of the Victorian Information Commissioner Office.

European Data Protection Board (2024) *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models.* Brussels; European Data Protection Board.

Datatilsynet, The Norwegian Data Protection Authority (January, 2018). *Artificial intelligence and privacy.* Oslo; The Norwegian Data Protection Authority.

Francis, S. (2024) *Navigating the Intersection of AI, Surveillance, and Privacy: A Global Perspective*. United Nations, Department of Economic and Social Affairs. Available at: https://sdgs.un.org/documents/francis-s-navigating-intersection-ai-surveillance-and-privacy-global-perspective-55502 [Accessed 21 February 2025].

Association of Southeast Asian Nations (2025), *AI Governance and Ethics – Generative AI.* Unknown*:* Association of Southeast Asian Nations.

European Group on Ethics in Science and New Technologies (2023) *Democracy in the Digital Age.* Brussels: European Commission.

Homo Digitalis (2024) *Πράξη για την τεχνητή νοημοσύνη – Ανάλυση διατάξεων για τις απαγορευμένες πρακτικές του άρθρου 5 του Κανονισμού 2024/1689*. Unknown: Homo Digitalis.

United Nations Human Rights Council (2024) *Mapping report: human rights and new and emerging digital technologies.* Geneva: United Nations Human Rights Council.

United Nations Human Rights Council (2022) *The right to privacy in the digital age.* Geneva: United Nations Human Rights Council.

European Commission (2024) *Annex to the Communication to the Commission: Approval of the content of the draft Communication from the Commission - Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)*. Brussels: European Commission.

ARTICLE 19 (2021). *Emotional Entanglement: China's Emotion Recognition Market and Its Implications for Human Rights*. London: ARTICLE 19.

EDPB and EDPS (2021) *Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Brussels: European Data Protection Board and European Data Protection Supervisor.

European Commission (2020) *White Paper on Artificial Intelligence - A European Approach to Excellence and Trust*. Brussels: European Commission.

Thesis

Kostopoulou, A. (2023). *Artificial Intelligence and Personal Data: Topical Issues on the Occasion of the EU AI Act*. Unpublished Master's Thesis. University of Piraeus,

4 THE AI ACT

4.1 The Artificial Intelligence Act - General Remarks

4.1.1 Introduction

While recognizing the transformative potential of AI and the economic, environmental and societal benefits across the entire spectrum of industries and social activities[13], governments also acknowledge its associated risks, including negative effects on employment, ethical values, privacy, intellectual property, and equality. (Cancella, 2024) In light of these developments, it is not surprising that legislators around the world are attempting to regulate AI, to protect against misuse without hindering innovation. In this context, on the 12th of July 2024 the Regulation (EU) 2024/1689 of the European Parliament and the Council was published in the Official Journal of the EU, laying down a harmonized, comprehensive legal framework to regulate the use of AI across its Member States; the Artificial Intelligence Act (AI Act). The Regulation entered into force on the 1st of August 2024, setting a global benchmark. This marks the first attempt by the EU to regulate AI, aspiring to create a new Brussels Effect. (Bird&Bird, 2024) During the following years, since the date of its application, AIA's provisions will be gradually applied[14].

It is important to highlight that the harmonized rules laid down in the AIA "should be without prejudice to existing Union law, in particular on data protection, consumer protection, fundamental rights, employment, and protection of workers, and product safety", as clearly mentioned in Recital 9.

4.1.2 Purpose

The purpose of the AIA, according to Recital 1, is firstly "to improve the functioning of the internal market by laying down a uniform legal framework in particular for the development, the

[13] Recital 4 AIA

[14] E.g. On February 2nd 2025, the prohibited practices ban applies (Chapter II). On August 2nd 2025, National authorities will be designated (Chapter III Section 4) and the obligations for General-Purpose AI (GPAI) (Chapter V) and governance (at EU and national level) will apply (Chapter VII).

placing on the market, the putting into service and the use of artificial intelligence systems (AI systems) in the Union, in accordance with Union values". Secondly, "to promote the uptake of human centric and trustworthy artificial intelligence (AI) while ensuring a high level of protection of health, safety, fundamental rights as enshrined in the Charter of Fundamental Rights of the European Union, including democracy, the rule of law and environmental protection, to protect against the harmful effects of AI systems in the Union, and to support innovation". And finally, to "ensure the free movement, cross-border, of AI-based goods and services, thus preventing Member States from imposing restrictions on the development, marketing and use of AI systems, unless explicitly authorised by this Regulation."

4.1.3 Legal Basis

The AIA was first introduced by the European Commission in April 2021 and agreed by the European Parliament and the Council in December 2023. It derives explicitly from the 2019 Ethics Guidelines on Trustworthy AI by the European Commission[15]. The legal basis for the Act, is in the first place Article 114 of the Treaty on the Functioning of the European Union (TFEU), "which provides for the adoption of measures to ensure the establishment and functioning of the internal market." (European Parliament and Council, 2021) As explained in Recital 3, some Member States had already considered national rules to ensure AI is trustworthy and respects fundamental rights. However, differing national regulations could fragment the internal market and reduce legal certainty. A consistent and high level of protection throughout the Union is deemed to be the most appropriate approach. Furthermore, to the extent that the AIA contains specific rules on the protection of individuals regarding the processing of personal data concerning restrictions of the use of AI systems in defined situations[16], Article 16 TFEU is regarded as a legal basis as well[17].

[15] Recital 27 AIA

[16] […]"restrictions of the use of AI systems for remote biometric identification for the purpose of law enforcement, of the use of AI systems for risk assessments of natural persons for the purpose of law enforcement and of the use of AI systems of biometric categorisation for the purpose of law enforcement,"[…]

[17] Recital 3 AIA

4.1.4 Scope

The AIA imposes specific obligations for operators, a term used to describe six categories of entities; providers, deployers, importers, distributors, product manufacturers and authorized representatives. Provider is the "natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge"[18] while deployer is "the 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".[19]

The Act sets rules for the placing on the market, the putting into service, and the use of AI systems and general purpose AI models and systems in the Union.[20] The regulated subject matter of the AIA therefore, are AI systems and general-purpose AI models[21] and systems.[22] The definition for the AI system, as mentioned above in chapter 2, is almost the same with the OECD's 2024 definition, except that any system with no capacity for 'adaptiveness after deployment' is apparently not AI for the AIA but can be according to OECD. (Greenleaf, 2024)

[18] Article 3(3) AIA
[19] Article 3(4) AIA
[20] Article 1(2) AIA

[21] 'general-purpose AI model' means 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. (Article 3(63) AIA)
[22] 'general-purpose AI system' means 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. (Article 3(66) AIA)

Regarding its territorial scope, while being a key piece of EU legislation, the AI Act also has extraterritorial effects. It applies when an AI system or general-purpose AI model is placed on the market or put into service in the Union, irrespective of whether those providers are established or located within the Union or in a third country[23], is used by deployers that have their place of establishment or are located within the Union [24], where the output produced by the AI system is used in the Union[25]. Explicitly excluded from the scope of the AIA are, amongst others, areas outside the scope of EU law (e.g. national security), AI systems which are not placed on the market or put into service in the Union, where the output is used in the Union exclusively for military, defense or national security purposes, regardless of the type of entity carrying out those activities or to AI systems or AI models, including their output, specifically developed and put into service for the sole purpose of scientific research and development. Article 2 AIA includes additional exemptions.

4.2 Risk Based Approach

4.2.1 The Approach

The AIA seeks to establish a proportionate and effective set of binding rules for AI systems, adopting a clearly defined, risk-based approach, categorizing systems from minimal to high risk. (Bird&Bird, 2024) Depending on the level of risk, AI systems have the respective requirements and the operators the respective obligations. According to the AIA definition risk means "the combination of the probability of an occurrence of harm and the severity of that harm".[26] The greater the risk to health, safety, or fundamental rights posed by a specific application, the more stringent the obligations. (Hacker, 2023) In order to determine the level of obligations for the operators, the risk category of the system must be established. (Greenleaf, 2024). The four risk-

[23] Article 2(1)(a) AIA
[24] Article 2(1)(b) AIA
[25] Article 2(1)(c) AIA

[26] Article 3(2) AIA

based categories are; prohibited, high risk, limited risk and minimal risk (the last two are not official terms).

4.2.2 Unacceptable risk

Chapter II Article 5 of the AIA establishes 8 prohibited AI practices which are considered to pose unacceptable risks, able to materially distort peoples' behavior or to threaten democratic societies, with great emphasis given on the biometric identification systems[27]. The concerned practices, according to Article 5, are in summary the following; a. Subliminal, manipulative, or deceptive techniques, b. Techniques able to exploit vulnerable groups by materially distorting their behavior and risking significant harm, c. Social scoring in certain cases, d. Criminality prediction based on profiling, e. Web or CCTV scraping for facial recognition databases, f. Inferences of emotions at workplaces or schools, g. Biometric categorization to infer sensitive data, h. Real-time remote biometric identification in public spaces for law enforcement purposes. These prohibitions are subject to exemptions and apply irrespective of the role of the operator. (Bird&Bird, 2024)

4.2.3 High Risk

The core of the proposed AI regulation, however, lies in the rules for the so-called high-risk AI systems mentioned in Article 6 AIA. (Hacker, 2023) An AI system falls within the scope of "high-risk" if under the following conditions; a. "the AI system is intended to be used[28] as a safety component of a product, or the AI system is itself a product[29]," listed in Annex I[30], b. "the product

[27] Article 5 AIA

[28] intended purpose is defined in article 3 paragraph 12 as: "the use for which an AI system is intended by the provider, including the specific context and conditions of use, as specified in the information supplied by the provider in the instructions for use, promotional or sales materials and statements, as well as in the technical documentation."

[29] Article 6(1)(a) AIA

[30] Eg. machinery, toys, recreational craft and personal watercraft, lifts/elevators, equipment and protective systems for potentially explosive atmospheres, radio equipment, pressure equipment, cableway

whose safety component pursuant to point (a) is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment, with a view to the placing on the market or the putting into service of that product" [31] listed in Annex I, and c. AI systems that "pose a significant risk of harm to the health, safety or fundamental rights of natural persons, including by materially influencing the outcome of decision making[32]" referred to in Annex III. AI systems identified as high-risk include AI technology used in;

-Management and operation of critical infrastructure (e.g. transport),

-Education and vocational training (e.g. scoring of exams),

-Safety components of products (e.g. AI application in robot-assisted surgery),

-Recruitment and HR (e.g. CV-sorting software for recruitment procedures),

-Essential private and public services (e.g. credit scoring systems denying citizens opportunity to obtain a loan),

-Crime analytics, Evidence gathering and evaluation (e.g. evaluation of evidence),

-Immigrant identification, migration risk and migration application assessment (e.g. automated examination of visa applications),

-Administration of justice and Democratic processes (e.g. AI solutions to search for court rulings),

-Remote biometric identification systems (e.g. Facial recognition technology). (Bird&Bird, 2024) (European Commission, 2024)

AI systems, taking into account their intended purpose as state of the art in AI, are subject to the strict requirements of Section 2 Articles 9-15 AIA[33];

- Adequate risk assessment and mitigation systems (Risk management system -Article 9),

installations, personal protective equipment, appliances burning gaseous fuels, medical devices, in vitro diagnostic medical devices, civil aviation, 2/3-wheel vehicles etc (Annex 1 AIA)

[31] Article 6(1)(b) AIA

[32] Article 6(2) and Article (3) AIA

[33] Article 8 AIA

- High quality of the datasets feeding the system to minimize risks and discriminatory outcomes (Data and data governance - Article 10),
- Detailed documentation and record keeping providing all information necessary on the system and its purpose for authorities to assess its compliance (Technical documentation - Article 11 and Record-keeping - Article 12)
- Clear and adequate information to the deployer (Transparency and provision of information to deployers - Article 13),
- Appropriate human oversight measures to minimize risk (Human oversight - Article 14) and
- High level of robustness, security and accuracy (Accuracy, robustness and cybersecurity - Article 15). (European Commission, 2024)

The obligations are further detailed for providers and other involved parties[34]. Providers of high-risk AI systems bear the most significant compliance responsibilities, like quality management[35], documentation[36] and conformity assessment[37][38]. Under the conformity assessment obligation, providers must demonstrate compliance with the requirements set out in Section 2, prior to being placed on the market or put into service[39]. Moreover, deployers of high-risk AI systems, shall perform an assessment of the impact on fundamental rights that the use of such system may have[40].

Providers of a general-purpose AI model that place such models on the market or integrate them with their own AI system and place them on the market or put them into service, are required to

[34] Requirements for providers of high-risk AI systems Articles 8-22, 43, 47-49, Requirements for deployers of high-risk AI systems Article 26, 27, Requirements for importers of high-risk AI systems Article 23, Requirements for distributors of high-risk AI systems Article 24, Requirement for third-party suppliers to high-risk systems Article 25
[35] Article 17 AIA
[36] Article 18 AIA
[37] Article 43 AIA
[38] Article 16 AIA
[39] Article 43 AIA
[40] Article 27 AIA

fulfill several transparency obligations. These obligations are directed both towards the AI Office and relevant authorities, as well as towards AI system providers who intend to integrate their systems with these general-purpose AI models[41]. In addition to the general obligations applicable to all general-purpose AI models' providers, the AI Act imposes further increased responsibilities on providers of general purpose AI models with systemic risk[42]. The obligations are and related to; models evaluation, assessment and mitigation of systemic risks, incident management and reporting, extended documentation and increased level of cybersecurity protection[43].

4.2.4 Limited and Minimal or No Risk

Limited risk systems that do "not pose a significant risk of harm to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making[44]", are mainly subject to transparency obligations[45]. It is required that providers or deployers of certain AI systems, limited and minimal risk, provide individuals with sufficient information by implementing appropriate disclaimers and marking[46]. Lastly, the Act allows the unrestricted use of minimal-risk AI, which includes the vast majority of AI systems currently in use in the EU fall into this category. (European Commission, 2024)

[41] Article 53, Article 54 AIA

[42] A general-purpose AI model is classified as a general-purpose AI model with systemic risk if it meets one of these two conditions: (a) it has "high impact capabilities" evaluated on the basis of technical tools and methodologies, or (b) is designated by the Commission as having capabilities or impact equivalent to those set out in point (Article 51(1), AIA)

[43] Article 55(1) AIA

[44] Article 6(3) AIA

[45] Article 6(4) AIA

[46] Article 50 AIA

4.3 Enforcement and Governance

4.3.1 Post-marketing Obligations

Apart from the ex ante requirements, the Act provides a multi-layered ex post enforcement framework, differentiating responsibilities based on the level of risk. Providers of high-risk AI systems are required to establish a post-market monitoring plan on how to actively and systematically analyze data to ensure compliance and address emerging risks[47]. Providers, and in some cases deployers, are obligated to report promptly serious incidents, such as harm to health, fundamental rights, or critical infrastructure disruptions[48].

4.3.2 Enforcement and Supervising Authorities

Each Member State must designate at least one market surveillance authority and responsible for overseeing compliance. These authorities can demand documentation, conduct inspections, and enforce corrective actions, including product recalls[49] and impose penalties[50]. Member States must also designate authorities to supervise AI systems' compliance with fundamental rights laws[51]. They must also designate a notifying authority to designate and supervise independent notified bodies.

On a European level, the implementation and governance of the Act is assigned to the European Artificial Intelligence Board, formed by representatives of the EU Member States, the Scientific Panel, formed by independent experts in the field of AI, and the Advisory Forum, composed by a diverse selection of stakeholders, both commercial and non-commercial. Regarding the supervision, investigation, enforcement and monitoring of providers of general-purpose AI models, the Commission will have exclusive responsibility. In practice, the implementation of

[47] Articles 72, 77(1) AIA

[48] Article 73 AIA
[49] Articles 70 and 74 AIA
[50] Articles 99-101 AIA
[51] Article 77, AIA

these tasks is delegated to the Commissions' AI Office[52], which shall take all the necessary actions to monitor their compliance with the Regulation[53]. Market surveillance authorities must collaborate with the AI Office to conduct compliance evaluations[54]. Lastly, for AI systems used by Union institutions, agencies, offices, and bodies, the European Data Protection Supervisor will be the supervisory authority[55].

Fines can be imposed by all the above mentioned authorities, reaching up to €35 million or 7% of annual global turnover for serious infringements[56].

4.3.3 Governance Framework

As per Recital 148, the successful implementation and enforcement of this Regulation require a governance framework that facilitates the coordination and development of central expertise at Union level. The collaborative governance structure provided by the Act integrates national and EU-level authorities with stakeholders in the field of AI to ensure effective coordination and support for its application[57]. The AI Office, the supranational body established by the Commission aims to "develop Union expertise and capabilities in the field of AI" with the facilitation of the Member States, while contributing to the implementation of the Act[58] and the drawing up of codes of practice[59]. Furthermore, the European Artificial Intelligence Board composed of representatives of the Member States with the participation of the European Data Protection Supervisor and the AI Office, shall advise and assist the Commission and the Member States in order to facilitate the consistent and effective application of this Regulation[60].

[52] Article 88 AIA
[53] Article 89 AIA
[54] Article 75 AIA
[55] Article 70(9) AIA
[56] Articles 99-101 AIA
[57] Recital 148 AIA
[58] Articles 64, 75 AIA
[59] Article 56 AIA
[60] Articles 65, 66s AIA

4.4 Final Remarks

The structure of the AIA reflects the EU's commitment to balancing innovation with harmonized regulatory oversight. It aims to increase trust in AI and ensure that this technology is used in a manner that respects the fundamental rights and safety of EU citizens. (Hacker, 2023) Overall, it is considered that the AIA "is a product safety law that regulates the safe technical development and use of AI systems and, with a few exceptions, it does not create any individual rights.' (Clark et al., 2024)

4.5 Citations for chapter 4

Articles

Cancela-Outeda C. (2024) 'The EU's AI act: A framework for collaborative governance, Internet of Things, Volume 27, Available at; https://doi.org/10.1016/j.iot.2024.101291. [Accessed 15.01.2024].

Greenleaf, G. (2024). 'EU AI Act: The 2nd most Important Data Privacy Law'. Privacy Laws & Business International Report, 189, 23. Available at; https://doi.org/10.2139/ssrn.4913686. [Accessed 15.01.2024].

Hacker, P. (2023). 'AI Regulation in Europe: From the AI Act to Future Regulatory Challenges'. Available at; https://doi.org/10.48550/arXiv.2310.04072 [Accessed 15.01.2024]

Legislation

European Parliament and Council Regulation (EU) 2024/1689 of the 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) (Text with EEA relevance) Available at; https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.

European Parliament and Council COM/2021/206 final Proposal for a Regulation of 21 April 2021 laying down harmonised rules on artificial intelligence and amending certain union legislative acts. Available at; https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52021PC0206 European Commission (2024). AI Act. Available at; https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai [Accessed 19.01.2024].

Websites

Clark, J., Demircan, M. and Kettas, K. (2024) '*Europe: The EU AI Act's relationship with data protection law: key takeaways. DLA Piper*, 25 April. Available at: https://privacymatters.dlapiper.com/2024/04/europe-the-eu-ai-acts-relationship-with-data-protection-law-key-takeaways/ [Accessed: 5 January 2025].

Bird&Bird (2024). European Union Artificial Intelligence (EU AI Act) Guide. Available at;

https://www.twobirds.com/en/insights/2024/global/european-union-artificial-intelligence-act-guide [Accessed 10.01.2024].

5 AI ACT AND PRIVACY: AIA'S INTERPLAY WITH DATA PROTECTION LAWS

5.1 Trustworthy and ethically sound AI

The goal of the AIA is to set a "proportionate and effective set of binding rules for AI systems" (Recital 26, AIA) with great emphasis in the aspect of trustworthiness and ethical soundness of the AI systems. In this regard, the Act is in alignment with the principles of the Ethics guidelines for trustworthy AI developed by the independent AI HLEG appointed by the Commission for trustworthy and ethically sound AI (2019); human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental well-being and accountability[61]. (Recital 27, AIA) Trustworthy AI consists of three key components, all of which should be upheld throughout the entire lifecycle of the system; lawfulness, "complying with all applicable laws and regulations", ethicalness, "ensuring adherence to ethical principles and values" and robustness, "both from a technical and social perspective since, even with good intentions, AI systems can cause unintentional harm", (Commission, 2019) These principles should be implemented, where feasible, in the design and use of AI models. In any case, they should form the foundation for developing codes of conduct under this Regulation. It seems like the Act has a strong intention

[61] "According to the guidelines of the AI HLEG, human agency and oversight means that AI systems are developed and used as a tool that serves people, respects human dignity and personal autonomy, and that is functioning in a way that can be appropriately controlled and overseen by humans. Technical robustness and safety means that AI systems are developed and used in a way that allows robustness in the case of problems and resilience against attempts to alter the use or performance of the AI system so as to allow unlawful use by third parties, and minimise unintended harm. Privacy and data governance means that AI systems are developed and used in accordance with privacy and data protection rules, while processing data that meets high standards in terms of quality and integrity. Transparency means that AI systems are developed and used in a way that allows appropriate traceability and explainability, while making humans aware that they communicate or interact with an AI system, as well as duly informing deployers of the capabilities and limitations of that AI system and affected persons about their rights. Diversity, non-discrimination and fairness means that AI systems are developed and used in a way that includes diverse actors and promotes equal access, gender equality and cultural diversity, while avoiding discriminatory impacts and unfair biases that are prohibited by Union or national law. Social and environmental well-being means that AI systems are developed and used in a sustainable and environmentally friendly manner as well as in a way to benefit all human beings, while monitoring and assessing the long-term impacts on the individual, society and democracy." (Rec. 27, AIA)

to safeguard fundamental rights including the right to privacy. According to Greanleef, the AIA positions itself as the "second most important privacy law" (2024).

5.2 AIA and other data protection laws

Article 2(7) states that "Union law on the protection of personal data, privacy and the confidentiality of communications applies to personal data processed in connection with the rights and obligations laid down in this Regulation", while Recital 69 states that "The right to privacy and to protection of personal data must be guaranteed throughout the entire lifecycle of the AI system. In this regard, the principles of data minimisation and data protection by design and by default, as set out in Union data protection law, are applicable when personal data are processed. Measures taken by providers to ensure compliance with those principles may include not only anonymisation and encryption, but also the use of technology that permits algorithms to be brought to the data and allow training of AI systems without the transmission between parties or copying of the raw or structured data themselves, without prejudice to the requirements on data governance provided for in this Regulation." "At the same time, the AI Act does not affect prohibitions that apply where an AI practice falls within other Union law. Thus, even where an AI system is not prohibited by the AI Act, its use could still be prohibited or unlawful based on other primary or secondary Union law (e.g., because of the failure to respect fundamental rights in a given case, such as the lack of a legal basis for the processing of personal data required under data protection law, discrimination prohibited by Union law, etc.)" (European Commission, 2024) Therefore, it is made clear that the provisions of the AIA are without prejudice to the GDPR and the LED. (Article 2(7), AIA) (Recital 10, AIA)

5.3 GDPR

As mentioned earlier, data protection is already regulated in the EU by the GDPR, the LED and the EUDPR. It is clear that there is a strong relation between the AIA and the GDPR. Many of the

data protection principles overlap with the principles and requirements for AI systems. Besides, the AIA is without prejudice to the GDPR, the latter being applicable when personal data processing takes place. (Recital 10, AIA) While AI is not explicitly mentioned in the GDPR, "many provisions in the GDPR are relevant to AI, and some are indeed challenged by the new ways of processing personal data that are enabled by AI". (European Parliamentary Research Service, 2020) "Article 22 regarding automated decision making, "serves as a form of indirect control over the use of AI systems, on the basis that AI systems are frequently used to take automated decisions that impact individuals". (Clark et al., 2024) On the other hand, as it will be further analyzed, in some cases tension and conflict appear between the two Regulations.

The GDPR sets out 7 principles for the processing of personal data; lawfulness, fairness and transparency (Article 5(1)(a)), purpose limitation (Article 5(1)(b)), data minimization (Article 5(1)(c)), accuracy (Article 5(1)(d)), storage limitation (Article 5(1)(e)), integrity and confidentiality (Article 5(1)(f)) and accountability (Article 5(2)). Moreover, Article 6 establishes 6 legal basis for data processing (consent, contract, legal obligation, vital interests, public interest, and legitimate interests), which remain applicable for AI systems that process personal data. According to Article 9 the processing of special categories of personal data (racial or ethnic origin, political opinions, religious or philosophical beliefs, trade union membership, genetic data, biometric data, health data, sex life, or sexual orientation) is prohibited unless an exception applies. High risk AI systems generally involve the processing of personal data and special categories of personal data. The providers and the deployers under the AIA, when they process personal data for the development or deployment of AI systems, will be also considered controllers under the GDPR (Article 24), depending on their role. (Clark et al., 2024)

Data subjects must be informed when their data is used for AI training and / or prediction. This includes the legal basis for such processing, general explanation of the logic and scope of the AI-system. In this context, individuals must always be able to exercise their right of restriction of processing (Article 18 GDPR and Article 20 EUDPR) as well as deletion / erasure of data (Article 16 GDPR and Article 19 EUDPR) Additionally, the controller should have explicit obligation to inform data subjects regarding the applicable periods for objection, restriction, deletion of data

etc. The AI system must be designed to fully comply with all data protection requirements through adequate technical and organizational measures. (EDPB and EDPS, 2021)

5.4 LED

The LED complements the GDPR and it concerns the processing of personal data by law enforcement authorities, which is out of scope of the GDPR, for the purposes of the prevention, investigation, detection and prosecution of criminal offences or the execution of criminal penalties (Recital 4, LED), creating a minimum standard of privacy protection across the EU. Some obligations for controllers are identical with the GDPR. Under LED the processing of personal data is lawful "only if and to the extent that processing is necessary for the performance of a task carried out by a competent authority" (Article 8, LED) while the processing of special categories of data is permitted only when it is strictly necessary.[62] Lastly, the rights granted to individuals are fewer compared to the GDPR. The LED does not apply to activities falling outside the scope of EU law, like national security. (Recital 14 and Article 2(3), LED)

Having established the key considerations, it is now crucial to assess whether the Act provides sufficient protection in addressing concerns regarding privacy and data protection in accordance with the existing data protection laws.

5.5 Citations for chapter 5

European Parliament and Council Regulation (EU) 2024/1689 of the of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No

[62] "(a) where authorised by Union or Member State law;
(b) to protect the vital interests of the data subject or of another natural person; or
(c)where such processing relates to data which are manifestly made public by the data subject." (Article 10, LED)

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) (Text with EEA relevance) Available at; https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) Available at; https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng.

Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA Available at; https://eur-lex.europa.eu/eli/dir/2016/680/oj/eng.

European Commission (2019) *Ethics guidelines for trustworthy AI*. Independent High-Level Expert Group on Artificial Intelligence. Brussels: European Commission

European Parliamentary Research Service (2020) *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*. Brussels: European Parliament.

European Commission (2024) *Annex to the Communication to the Commission: Approval of the content of the draft Communication from the Commission - Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)*. Brussels: European Commission.

EDPB and EDPS (2021) *Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Brussels: European Data Protection Board and European Data Protection Supervisor.

Greenleaf, G. (2024). 'EU AI Act: The 2nd most Important Data Privacy Law'. Privacy Laws & Business International Report, 189, 23. Available at; https://doi.org/10.2139/ssrn.4913686. [Accessed 15.03.2025].

Clark, J., Demircan, M. and Kettas, K. (2024) '*Europe: The EU AI Act's relationship with data protection law: key takeaways. DLA Piper*, 25 April. Available at: https://privacymatters.dlapiper.com/2024/04/europe-the-eu-ai-acts-relationship-with-data-protection-law-key-takeaways/ [Accessed: 11 April 2025].

6 AI ACT PROHIBITED PRACTICES

6.1 Prohibited practices

The AIA aims to prevent the misuse of AI for "manipulative, exploitative and social control practices". As mentioned earlier, the AIA, in Article 5, establishes 8 prohibited practices which pose a significant threat to fundamental rights and interfere with the right to privacy, amongst other rights. The AI practices declared as prohibited can have severe effects on informational, decisional, physical and behavioral privacy, as mentioned in the previous chapter. The Act considers the risks of AI to privacy and draws strict rules. As defined by the Act, the Commission published[63] the "Guidelines on prohibited AI practices" aiming to address the unacceptable AI practices, increase legal clarity, interpret the prohibitions, provide examples and ensure the effective and consistent application of the AIA across the EU. (European Commission, 2025)

Although it is important to notice that while these Guidelines provide valuable insights, they are non-binding. Any authoritative interpretation of the AI Act may ultimately only be given by the Court of Justice of the European Union (CJEU). (European Commission, 2025) They serve as a "practical guidance to assist competent authorities under the AI Act in their enforcement activities, as well as providers and deployers of AI systems in ensuring compliance with their obligations under the AI Act." (European Commission, 2025)

It is also important to note that even though a practice is not prohibited by the AIA it may be prohibited by other EU legislation, in this case EU data protection laws. (European Commission, 2025)

[63] the Commission has approved the draft guidelines, but not yet formally adopted them. (European Commission, 2025)

6.2 Manipulative or deceptive techniques - Article 5(1)(a) AIA

According to Article 5(1)(a), the AIA prohibits the "the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective, or the effect of materially distorting the behavior of a person or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause that person, another person or group of persons significant harm".

As further analyzed in recital 29, "Such AI systems deploy subliminal components such as audio, image, video stimuli that persons cannot perceive, as those stimuli are beyond human perception, or other manipulative or deceptive techniques that subvert or impair a person's autonomy, decision-making or free choice in ways that people are not consciously aware of those techniques or, where they are aware of them, can still be deceived or are not able to control or resist them". Some examples mentioned in recital 29 are machine-brain interfaces or virtual reality, used in various fields, amongst them in manipulative marketing strategies, (Homo Digitalis, 2024) which may materially distort a person's behaviour in a significantly harmful manner. This prohibition applies to both providers and deployers of AI systems and a plausible link between the AI system and the harm is required. Nevertheless, it is not necessary for the provider or the deployer to have the intention of causing significant harm. (Recital 29, AIA)

The European Commission specifies the terms subliminal, purposefully manipulative or deceptive techniques, since they are not defined in the AIA. Firstly, subliminal techniques, using stimuli delivered through audio, visual, or tactile media that are too brief or subtle for people to notice, "must be capable of influencing behavior in ways in which the person remains unaware of such influence, how it works, or its effects on the person's decision-making or value and opinion formation" (European Commission, 2025) Some examples of subliminal techniques that may be prohibited are visual subliminal messages, auditory subliminal messages, subvisual and subaudible cueing, embedded Images etc.

Secondly, purposefully manipulative techniques "should be understood as techniques that are designed or objectively aim to influence, alter, or control an individual's behaviour in a manner that undermines their individual autonomy and free choices", typically designed to "exploit cognitive biases, psychological vulnerabilities, or situational factors that make individuals more susceptible to influence" (European Commission, 2025) Recital 29 AIA makes clear that the prohibition in Article 5(1)(a) also covers techniques where individuals, even if they are aware of the influence attempt, they may still be unable to control or resist its manipulative effects. One example provided is AI systems that create and tailor highly persuasive messages based on an individual's personal data influencing their behaviour or choices to a point of creating significant harm. AI systems that deploy such techniques, regardless of whether the provider or the deployer has designed or used the system in this way, for example if the system has learned manipulative techniques due to the fact that the training data contain many instances of manipulative techniques, still fall under the prohibition as well. (European Commission, 2025)

Thirdly, deceptive techniques "deployed by AI systems should be understood to involve presenting false or misleading information with the objective or the effect of deceiving individuals and influencing their behaviour in a manner that undermines their autonomy, decision-making and free choices." (European Commission, 2025) Like with manipulative techniques, it covers AI systems that use deceptive techniques without the provider's or deployer's intention. This prohibition is closely related to the deployer's obligation to label 'deep fakes' and certain AI generated text publications on matters of public interest stated in Article 50(4) AIA, as well as the provider's obligation to ensure AI systems interacting with people are designed in a way to clearly inform people that they are engaging with AI instead of a human. On the other hand, the prohibition of deceptive techniques has a more limited scope. For example it may apply to cases where "a chatbot or deceptive AI-generated content presents false or misleading information in ways that aim to or have the effect of deceiving individuals and distorting their behaviour that would not have happened if they were not exposed to the interaction with the AI system or the deceptive AI generated content, in particular if this has not been visibly disclosed". Both manipulative and deceptive techniques are not prohibited when they are incidental (European Commission, 2025).

The prohibition applies when significant harm is caused due to material distortion of a person's behavior, "in particular having sufficiently important adverse impacts on physical, psychological health or financial interests", regardless of the intention of harm. (Recital 29, AIA) The Commission clarifies that a substantial impact on a person's behavior is required, involving a degree of coercion, manipulation, or deception that exceeds lawful persuasion, which falls outside the scope of the prohibition, such that their autonomy and free choices are undermined. (European Commission, 2025)

Since it is not defined in the AIA, for the interpretation of "material distortion of behavior" the Commission suggests to follow the Directive 2005/29/EC (Unfair Commercial Practices Directive or 'UCPD').[64] Like under the UCPD, "market surveillance authorities must also investigate each case's specific facts and circumstances, assessing whether the subliminal, purposefully manipulative or deceptive technique deployed by the AI system is likely to appreciably impair the decision-making, individual autonomy and free choice of an 'average' individual within a targeted group when the system affects a group of persons in a manner that is reasonably likely to cause significant harm", therefore judging each case in concreto. (European Commission, 2025) In consistency with consumer protection law, it suffices that the harm is reasonably likely to occur, without the need to have fully materialized. (European Commission, 2025)

According to Article 5(1)(a), the main types of harm include physical, psychological and financial harm of persons or group of persons and the guidelines further analyze the terms. Physical harm includes any injury or damage to a person's life, health and material damage to property, with serious and irreversible consequences. (European Commission, 2025) Psychological harm refers to negative impacts on a person's mental health, as well as their emotional and psychological

[64] "The UCPD prohibits various unfair, misleading, and aggressive commercial practices (Articles 5 to 9 UCPD) capable of causing consumers to make transactional decisions that they would otherwise not have made. According to the CJEU and the Commission guidance on the UCPD, there is no need to prove that a consumer's economic behaviour has been distorted, it suffices to establish that a commercial practice is 'likely' (i.e., capable) of impacting an average consumer's transactional decision.The CJEU has also underscored that even accurate information may be misleading if presented in a way that distorts the consumer's decision-making process. National enforcement authorities are tasked to investigate the specific facts and circumstances of each case (in concreto) and to evaluate the potential impact of the practice on the average consumer's decision-making process (in abstract)." (European Commission, 2025)

well-being, for example "suicidal behaviours and risks of harming other persons". (European Commission, 2025) Nevertheless a case by case assessment is required to determine the seriousness. Financial and economic harm may encompass financial loss, financial exclusion and economic instability. In assessing the harm, it is important to have in mind that they may be combined, for example harm by fostering addictive behaviours, anxiety, and depression that may subsequently result in physical harm, such as insomnia and other stress-related health issues and physical problems. (European Commission, 2025)

In order for the prohibition to apply, the harm must be significant, implying significant adverse impacts. "The determination of 'significant harm' is fact -specific, requiring careful consideration of each case's individual circumstances and a case-by-case assessment, but the individual effects should be always material and significant in each case." (European Commission, 2025) Some indicators to assess 'significant harm' include the severity of harm, context and cumulative effects, scale and intensity, affected persons' vulnerability and duration and reversibility. Nevertheless, there is a number of cases where the threshold of significant harm is not reached even through individuals are harmed by AI systems deploying subliminal, purposefully manipulative, or deceptive techniques. (European Commission, 2025)

Lastly, a casual link between the practice and the harm (occurred or likely to occur) is necessary for the prohibition to apply. "Factors external to the AI system which are outside the control of the provider or the deployer" are not sufficient to establish a casual link. (Recital 29, AIA) According to the guidelines, it is necessary that the provider and the deployer could have reasonably foreseen the significant harm according to universally accepted criteria and industry standards, opacity and transparency requirements, respect to individual autonomy and compliance with relevant applicable legislation. (European Commission, 2025)

Even though the Commission Guidelines aim to clarify this provision and provide examples, there are still some issues that must be addressed. First of all, the prohibition depends on the occurrence (or likely occurrence) of significant harm and not on the mere deployment of subliminal, purposefully manipulative or deceptive techniques or even the mere material distortion of behavior due to these techniques. (Homo Digitalis, 2024) Prior to the potential occurrence of harm, the right to privacy has already been compromised. These techniques

interfere with personal autonomy and violate individuals behavioral and decisional privacy by distorting their thinking process and acts. Secondly, the notion of significant harm is still vague and is judged case by case since the individual effects differ. The guidelines provide some criteria to identify significant harm but still the measurement and method of evaluation is unclear. The measurement of the material distortion of behavior also appears vague, since it depends on the individual's inner state. (Homo Digitalis, 2024) Therefore, there may be numerous AI systems that deploy subliminal, purposefully manipulative, or deceptive techniques that materially distort behavior and cause a level of harm that will likely not reach the threshold of significant. Also, given the nature of AI systems and the vagueness of significant harm, the intensity of the harm is difficult to be foreseen, taking into consideration AI's "black box" problem and adaptability capabilities. It is important to note that apart from physical, psychological and financial, the harm may be directed towards fundamental rights and the rule of law. (Homo Digitalis, 2024)

Lastly, according to recital 29, "common and legitimate commercial practices, for example in the field of advertising, that comply with the applicable law should not, in themselves, be regarded as constituting harmful manipulative AI-enabled practices". There is a very fine line between manipulative and highly persuasive lawful techniques that cause distortion of behavior. Especially in the field of advertising, since AI enhanced advertising in lawful, a very careful examination on a case to case basis is needed to ensure protection of consumer autonomy and decision making from manipulation. After all, manipulation can occur even if persons are aware of these techniques and the influence at play. (Recital 29, AIA)

6.3 Harmful exploitation of vulnerabilities - Article 5(1)(b) AIA

According to Article 5(1)(b), the Act prohibits "the placing on the market, the putting into service or the use of an AI system that exploits any of the vulnerabilities of a natural person or a specific group of persons due to their age, disability or a specific social or economic situation, with the objective, or the effect, of materially distorting the behaviour of that person or a person belonging

to that group in a manner that causes or is reasonably likely to cause that person or another person significant harm". Specific group of persons is interpreted within the meaning of Directive (EU) 2019/882[65] and socioeconomic situations are those "likely to make those persons more vulnerable to exploitation such as persons living in extreme poverty, ethnic or religious minorities." (Recital 29, AIA)

This provision is an extension of Article 5(1)(a) and concerns vulnerable individuals or groups that, due to certain characteristics, are more susceptible to manipulation and exploitation and their ability to make an informed decision can be affected more easily by making use of such vulnerabilities. While "any vulnerabilities" can be understood to encompass a broad spectrum of categories more susceptible to influence, the Act limits the prohibition to vulnerability due to age, disability and socioeconomic situation. (European Commission, 2025) This means, children and older people, people with "long-term physical, mental, intellectual, and sensory impairments" and persons living in extreme poverty, ethnic or religious minorities, or other relatively stable and long-term characteristics. On the other hand, biased systems that disproportionately target socio-disadvantaged individuals are not automatically considered exploitative. The concepts of material distortion of behavior and significant harm are the same with Article 5(1)(a), except the requirement to "appreciably impair the ability to make an informed decision". (European Commission, 2025)

The same problematics appear regarding the threshold of significant harm and lawful persuasion as mentioned above. The mere exploitation of vulnerabilities and behavior distortion are not enough to prohibit an AI practice. Also, even though there are plenty of instances where an individual or group are more susceptible to influence, "it follows from the wording of Article 5(1)(b) AI Act that this susceptibility must be the result of the person belonging to one of the groups ('due to')". (European Commission, 2025)

[65] Directive (EU) 2019/882 of the European Parliament and of the Council of 17 April 2019 on the accessibility requirements for products and services (OJ L 151, 7.6.2019, p. 70).

6.4 Social scoring - Article 5(1)(c) AIA

According to Article 5 AIA, "the placing on the market, the putting into service or the use of AI systems for the evaluation or classification of natural persons or groups of persons over a certain period of time based on their social behaviour or known, inferred or predicted personal or personality characteristics, with the social score leading to either or both of the following:

(i) detrimental or unfavourable treatment of certain natural persons or groups of persons in social contexts that are unrelated to the contexts in which the data was originally generated or collected;

(ii) (ii) detrimental or unfavourable treatment of certain natural persons or groups of persons that is unjustified or disproportionate to their social behaviour or its gravity" shall be prohibited.

The prohibition covers the placing on the market, the putting into service or the use of the AI systems by the public and private sector (Recital 31, AIA) that evaluate or classify individuals and produce a score "in various forms, such as a mathematical number (for example, from 0 to 1), a ranking, or a label. (European Commission, 2025) "While 'evaluation' suggests the involvement of some form of an assessment or judgement about a person or group of persons, a simple classification of persons or groups of persons based on characteristics, such as their age, sex, and height, need not necessarily lead to evaluation. (European Commission, 2025) Although not explicitly stated by the Act, the 'evaluation' also refers to the concept of 'profiling' as regulated by the GDPR. Therefore, "Profiling of natural persons under EU data protection law, when conducted through AI systems, may therefore also be covered by Article 5(1)(c) AI Act." (European Commission, 2025)

Furthermore, the prohibition requires that the evaluation or classification should rely on data collected over a "certain period of time", suggesting that the assessment should not be confined to a one-time rating or grading based on data or behavior from a specific, isolated individual context (European Commission, 2025) and "based on their social behaviour or known, inferred or predicted personal or personality characteristics". Social behaviour means actions, behaviour, habits, interactions within society, etc. in private (e.g. volunteering), business contexts (e.g.

payment of debts) and relations with public and private entities, government, police, and the law (e.g. obeying traffic rules) and personal or personality characteristics means a variety of characteristics like sex, income, family situation, health, type of car etc. (European Commission, 2025)

The prohibition applies when the scoring created by or with the assistance of an AI system must lead, or is intended or capable of leading, to a detrimental or unfavourable treatment, including cases where human assessment is also involved. (European Commission, 2025) 'Unfavourable treatment' means that "as a result from the scoring, the person or group of persons must be treated less favourably compared to others without necessarily requiring a particular harm or damage while 'detrimental' treatment "requires the person or group of persons to suffer certain harm and detriment from the treatment". The result of unfavourable or detrimental treatment may be discrimination, in a sense that goes beyond EU non-discrimination law, which applies only to certain protected groups (e.g., age, ethnic and racial origin, sex, religion). (European Commission, 2025)

Under Article 5(1)(c)(i), detrimental or unfavourable treatment must take place in social contexts not related to the contexts in which the data was originally generated or collected. (European Commission, 2025) "This implies not only that the persons may be treated in an unfavourable or detrimental manner due to the social score, but also that the data about their social behaviour or their known, inferred or predicted personal or personality characteristics are generated or collected in social contexts unrelated to the one in which the scoring takes place". (European Commission, 2025) In most cases, this happens contrary to the reasonable expectations of the individuals involved and in breach of Union data protection law (especially the principle of purpose limitation as expressed in Article 5 (1) (b) GDPR) and possibly other applicable rules that specify the types of data and sources considered relevant and necessary for the evaluation or classification. (European Commission, 2025) In the (ii) scenario of Article 5(1)(c), the prohibition applies when the treatment resulting from the scoring is unjustified or disproportionate to the gravity of the social behaviour. A case-by-case assessment is required to examine the impact of the interference to a person's fundamental rights compared to the gravity of their social behavior, taking into account the legitimate aim pursued and the principle of proportionality (European

Commission, 2025). A combination of these results is possible, for example an AI-based governmental system that monitors and rates individuals based on their behaviour in different aspects of their social lives and restricts people with lower scores from public services, traveling, finding a job etc. "The system leads to excessive surveillance of individuals and detrimental treatment in contexts unrelated to the social behaviour used to determine the social score (e.g. job opportunities are influenced by social media activity), while also imposing excessive penalties for minor infractions (e.g. social and financial disadvantages for relatively minor offences)". (European Commission, 2025)

The prohibition of Article 5(1)(c) "should not affect lawful evaluation practices of natural persons that are carried out for a specific purpose in accordance with Union and national law. (Recital 31, AIA) Therefore, the scoring of natural persons is not per se prohibited if it's in compliance with Union law. According to the guidelines, credit scoring, risk scoring, and underwriting are crucial components of the services provided by financial and insurance companies. (European Commission, 2025) "In other words, AI systems which evaluate or classify individuals for the purposes of generating a social score in a lawful manner and for a specific purpose in the related context as that in which the personal data used for the score were collected are not prohibited, provided that any detrimental or unfavourable treatment from using the score is justified and proportionate to the gravity of the social behaviour." (European Commission, 2025) Compliance with Union law, like data protection laws, ensures that the treatment is justified and proportionate and that the collection and processing of the data is lawful. (European Commission, 2025)

This prohibition aims to safeguard the right to dignity and non-discrimination, the values of equality and justice (Recital 31, AIA) and the right to privacy as well. (European Commission, 2025) The guidelines specifically state that the terms social behaviour and personal and personality characteristics are very broad and encompass wide range of information. On the other hand, lawful evaluation practices fall outside the scope of the prohibition and will be considered high risk systems. (Jakubowska et al, 2024) In contrast with human decision making, automated decisions focus on optimizing performance and predicting outcomes, with decisions relying on training data and the system's architecture. Lack of transparency, a problem inherent with the

nature of AI systems, makes the evaluation of the systems difficult. The translation of human behavior into data (tokenization) and their process by AI systems still appears problematic. Automated decision making can lead to subjective and unreliable evaluation since it's a mechanism that reduces complex and multidimensional decisions into simple decisions based on scores, under the illusion that these are fairer and more unbiased because they are based on data. (Homo Digitalis, 2024) Also such systems falsely follow the logic that the more data collected the more precise the results will be.

In any case, social scoring may violate data protection laws. A tension exists between AI systems that perform social scoring and the purpose limitation principle of Article 5(1)(b) GDPR and the data minimization principle of Article 5(1)(c) GDPR, since these systems often gather extensive data about individuals and re-use them for purposes further than originally collected, as well as the principles of accuracy Article 5(1)(d) GDPR and storage limitation Article 5(1)(e) GDPR. (European Parliamentary Research Service, 2020) The GDPR provisions regarding the legal ground for processing (Article 6 GDPR) and the rules of fully automated decision making (Article 22 GDPR) must also be complied with. Thus, under the GDPR the processing of personal data in violation of these principles is unlawful, regardless of whether detrimental or unfavorable treatment of an individual based on social scoring occurs. The EDPB and EDPS claim that any type of social scoring should be prohibited and add that that in the context of law enforcement, Article 4 LED already imposes significant limits, if not in practice prohibits such type of activities. (2021)

6.5 Individual risk assessment and prediction of criminal offences - Article 5(1)(d) AIA

Article 5(1)(d) AIA prohibits "the placing on the market, the putting into service for this specific purpose, or the use of an AI system for making risk assessments of natural persons in order to assess or predict the risk of a natural person committing a criminal offence, based solely on the profiling of a natural person or on assessing their personality traits and characteristics; this

prohibition shall not apply to AI systems used to support the human assessment of the involvement of a person in a criminal activity, which is already based on objective and verifiable facts directly linked to a criminal activity". This prohibition is in line with the presumption of innocence and with the principle that "natural persons in the Union should always be judged on their actual behaviour". (Recital 42, AIA) "Article 5(1)(d) AI Act is without prejudice to Article 11(3) LED, which prohibits profiling resulting in (direct or indirect) discrimination". (European Commission, 2025)

Crime prediction AI systems are designed to identify patterns with historical data and generate risk scores for the likelihood of a crime occurring and are mainly used by law enforcement authorities. The risk assessments and predictions concern in principle future criminal activities or crimes that are considered a potential risk of being committed at the moment. (European Commission, 2025) Article 5(1)(d) AIA does not prohibit crime prediction and risk assessment practices as such. It only applies to cases where risk assessments are made solely a) on the profiling of a natural person or b) assessing a natural person's personality traits and characteristics, or both. Private actors may also fall within the scope of the prohibition when "they are entrusted by law to exercise public authority and public powers for the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties" or when they assess or predict the risk of a person committing a specific crime (e.g. terrorism financing) is necessary for compliance with a legal obligation. (European Commission, 2025)

Article 5(1)(d) explicitly uses the term 'profiling', in the same meaning as in Article 4(4) GDPR (Article 3(52), AIA) and Article 3(4) LED as 'any form of automated processing of personal data consisting of the use of personal data to evaluate certain personal aspects relating to a natural person, in particular to analyse or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behaviour, location or movements'. In the context of Article 5(1)(d) the purpose of profiling is predicting criminal activity. (European Commission, 2025) If profiling as defined in Article 4(4) GDPR cannot be established, the prediction of a criminal offence based on the assessment of personality traits and characteristics is also prohibited, which includes a broad category of characteristics related to an individual, as analyzed in Article 5(1)(c), (European Commission, 2025) for example

nationality, place of birth, place of residence, number of children, level of debt or type of car. (Recital 42, AIA)

The prohibition only applies in cases where the risk assessment is based 'solely' on profiling or assessing personality traits and characteristics. This condition leaves out of scope of the prohibition numerous situations. Article 5(1)(d) in its last phrase states that "the prohibition does not apply to AI systems used to support the human assessment of the involvement of a person in a criminal activity, which is already based on objective and verifiable facts directly linked to a criminal activity". Recital 42 points out that "in this context one should think, in particular but not necessarily exclusively, of a situation in which a reasonable suspicion in respect of the natural person concerned already exists". (European Commission, 2025) To prevent the circumvention of the prohibition and ensure its effectiveness, any such additional elements must be genuine, significant, and meaningful in order to justify the conclusion that the prohibition does not apply. (European Commission, 2025) This situation "is not necessarily the only one in which the prohibition does not apply". (European Commission, 2025) Nevertheless, the criteria are still vague and leave room for misinterpretation. Even if used in an assistive manner the systems may produce inaccurate results and "may perpetuate or even reinforce biases, and result in crucial individual circumstances being 'overlooked' when these circumstances are not part of the data set or considered in the algorithms on which the particular AI system operates". (European Commission, 2025)

Where a system falls within the scope of the exclusion of the prohibition, it will be considered a high risk system if intended to be used by law enforcement authorities or on their behalf, and be subject to appropriate requirements and safeguards (Annex III, point 6(d), AIA). Taking into consideration that the use of AI systems by law enforcement authorities is characterized by "a significant degree of power imbalance and may lead to surveillance, arrest or deprivation of a natural person's liberty as well as other adverse impacts on fundamental rights guaranteed in the Charter" Recital 59 highlights the importance of the high quality of the data to ensure accurate and robust results and avoid discriminatory outcomes.

Furthermore, "the use of AI systems for making risk assessments in the context of private entities' ordinary course of business and with the aim of protecting their own private interests, whilst the

fact that those risk assessments may relate to the risk of criminal offences being committed merely as a purely accidental and secondary circumstance, is not deemed to be covered by the prohibition". (European Commission, 2025)

Also excluded from the scope of the prohibition are location and event-based predictions and apparently the prohibition only covers a part of predictive policing. (Access Now, 2024) It is not always clear how to distinguish location or event based from person based but "To the extent that an AI system carries out location-based predictive policing and then considers the risk score of the location as an aspect in the profiling of a person, that system should be considered person-based and in principle covered by Article 5(1)(d) AI Act, although it may fall outside the scope of the prohibition on other grounds" (European Commission, 2025) In any case, the EDPB and EDPS recognize the detrimental effect of predictive policing on human dignity and the objectification of human beings and propose a total prohibition of such practice. (2021)

6.6 Untargeted scraping of facial images - Article 5(1)(e) AIA

Next, Article 5(1)(e) AIA prohibits "the placing on the market, the putting into service for this specific purpose, or the use of AI systems that create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage". Indeed, "the untargeted scraping of facial images from the internet and from CCTV footage seriously interferes with individuals' rights to privacy and data protection and deny those individuals the right to remain anonymous", (European Commission, 2025) since facial images, in comparison with other biometric data, provide the most effective basis for remote identification.

A facial recognition database can match a human face from a digital image or video frame with a database of faces, comparing it to the images stored and assessing whether there is a probable match. It is enough that the database can be utilized for facial recognition, it is not necessary for the database's sole purpose to be facial recognition. (European Commission, 2025) This practice can be utilized for example for post remote biometric identification. The prohibition only applies to AI systems that do untargeted scraping, meaning "without a specific focus on a given

individual or group of individuals", gathering as much information and data as possible from the internet like social media platforms or from CCTV footage in places such as parks, airports streets etc., without respecting the opt-out of internet protocols. (European Commission, 2025) The guidelines highlight that the prohibition should be interpreted in a way that prevents the circumvention of the prohibition and considers prohibited practices where "the end-result is functionally the same as pursuing untargeted scraping from the outset", for example a system that uses targeted scraping selecting specific groups of individuals or other criteria each time and then combining them. (European Commission, 2025) In addition, the images are collected with any associated information "such as the source of the image (URL), the geo-localisation, and sometimes the names of the individuals". It is important to note that the correct interpretation of the facial image database must include the storage of only biographical information or URLs and not the actual facial images, otherwise these facial recognition systems would fall outside the scope of the prohibition. (Access Now, 2024) Lastly, targeted data scraping is not adequately determined and still leaves room for scraping of facial images of persons who are unrelated to the target of the scraping, violating their privacy. This is particularly important in situations of imbalance of power. (Homo Digitalis, 2024) Data protection laws shall be applied accordingly in any case.

6.7 Emotion recognition - Article 5(1)(f) AIA

Another prohibited practice is "the placing on the market, the putting into service for this specific purpose, or the use of AI systems to infer emotions of a natural person in the areas of workplace and education institutions, except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons". Recital 44 mentions that the rational behind the prohibition is the lack of reliability of such systems and the serious concerns about their scientific basis particularly because the expression of emotions differs across cultures, situations and

individuals. Plus, they might lead to discriminatory outcomes and violate the rights and freedoms of natural persons, including the right to privacy.

Recital 44 clarifies that the prohibition applies to systems able to infer and identify emotions or intentions based on a person's biometric data. Physical states like fatigue or pain are excluded. Article 3(34) AIA defines 'biometric data' as 'personal data resulting from specific technical processing relating to the physical, physiological or behavioural characteristics of a natural person, such as facial images or dactyloscopic data. "Behavioural biometrics monitor the distinctive characteristics of movements, gestures, and motor-skills of individuals as they perform a task or series of tasks" (European Commission, 2025)

This provision is limited to "areas of workplace and education institutions" except where it is used for medical or safety reasons. The guidelines clarify that this does not include general aspects of wellbeing, such as stress level. This exception must be interpreted narrowly and not be used as a justification for implementing such systems. "For example, an AI system intended to detect burnout or depression at the workplace or in education institutions would not be covered by the exception and would remain prohibited." (European Commission, 2025) On the other hand deployers are responsible to employ adequate safeguards "however, it cannot be completely avoided that such systems also infer the emotions of those employees" (European Commission, 2025)

Other emotion recognition systems that do not fall under the prohibition will be considered high risk AI systems and at the same time some might be prohibited under Article 5(1)(a) and (b) AI Act or data protection laws. Nevertheless, it is problematic that the AIA limits the prohibition to areas of workplace and education institutions "but are still allowed when used by law enforcement and migration authorities". This signals "the EU's will to test the most abusive and intrusive surveillance systems against the most marginalised in society". (Jakubowska et al., 2024) In addition, by contrast it can be concluded that AI emotion recognition systems are allowed in many other cases like analyzing consumer behavior, targeted advertising or law enforcement. 'Crowd control' systems, which are generally systems that enable "the control and monitoring of the behaviour of groups to maintain (public) order and event safety" without inferring emotions of concrete natural persons, are also excluded from the prohibition. (European Commission, 2025)

This provision does not adequately protect the individual's fundamental right to privacy from intrusive systems, not even in environments where there is an imbalance of power. It is pointed out that in spite of the AIA's limited scope, emotion recognition based on biometric data is covered by the GDPR. So, the assessment of whether an AI system that processes personal data for emotion recognition is allowed or not must be based on the GDPR and its provisions regarding the processing of special categories of personal data (Article 9 GDPR). (EDPS, 2021) The EDPB and EDPS consider that, except for very specific cases (namely for health or research) "the use of AI to infer emotions of a natural person is highly undesirable and should be prohibited". (2021)

6.8 Biometric categorization for certain 'sensitive' characteristics - Article 5(1)(g) AIA

According to Article 5(1)(g) the following practice shall also be prohibited; "the placing on the market, the putting into service for this specific purpose, or the use of biometric categorisation systems that categorise individually natural persons based on their biometric data to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation; this prohibition does not cover any labelling or filtering of lawfully acquired biometric datasets, such as images, based on biometric data or categorising of biometric data in the area of law enforcement". The rationale behind the provision is the wide variety of information, including sensitive information that can be extracted or inferred from biometric information, even without concerned individuals knowing so, which could lead to discriminatory and unfair treatment. (European Commission, 2025)

The prohibition applies to systems that perform biometric categorization, meaning "the process of establishing whether the biometric data of an individual belong to a group with some pre-defined characteristic." It's about "assigning an individual to a certain category", not identifying or verifying their identity. (European Commission, 2025) For example, identifying the sexual orientation of a person based on their photo. (Homo Digitalis, 2024)

Features "intrinsically linked to another commercial service, meaning that the feature cannot, for objective technical reasons, be used without the principal service", fall out of the scope of Article 5(1)(g), without meaning to circumvent the applicability of the rules of the AIA. (European Commission, 2025) (Recital 16, AIA)

For the prohibition to apply, the biometric categorization systems must have as an objective to deduce the certain characteristics mentioned Article 5(1)(g) AIA; race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation. Therefore, AI systems that can deduce other characteristics, for example personality or social behaviour, are not prohibited. Group categorization, instead of individual, should not be used as a justification to reach the same results either. (European Commission, 2025) Moreover, the prohibition does not cover the "lawful labelling, filtering or categorisation of biometric data sets acquired in line with Union or national law according to biometric data, such as the sorting of images according to hair colour or eye colour, which can for example be used in the area of law enforcement either".

"AI systems intended to be used for biometric categorisation according to sensitive attributes or characteristics protected under Article 9(1) GDPR on the basis of biometric data, in so far as these are not prohibited under this Regulation, are classified as highrisk under the AIA". The GDPR and the LED also apply accordingly. (European Commission, 2025)

6.9 Real-time remote biometric identification (RBI) for law enforcement purposes - Article 5(1)(g) AIA

The last prohibition regards real time remote biometric identification (RBI) in publicly accessible spaces for purposes of law enforcement and provides 3 situations in which the use may be permitted. More specifically "The following AI practices shall be prohibited: h) the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purpose of law enforcement, unless and in so far as such use is strictly necessary for one of the following objectives:

i) the targeted search for specific victims of abduction, trafficking in human beings or sexual exploitation of human beings, as well as the search for missing persons;

ii) ii) the prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons or a genuine and present or genuine and foreseeable threat of a terrorist attack;

iii) iii) the localisation or identification of a person suspected of having committed a criminal offence, for the purpose of conducting a criminal investigation or prosecution or executing a criminal penalty for offences referred to in Annex II and punishable in the Member State concerned by a custodial sentence or a detention order for a maximum period of at least four years.

Point (h) of the first subparagraph is without prejudice to Article 9 of Regulation (EU) 2016/679 for the processing of biometric data for purposes other than law enforcement."

Article 5(5) AIA states that Member States shall decide which of the three situations will allow in their territory, for which the rules for high-risk AI systems will apply (Article 6(2) and point a) of Annex III AI Act). The placing on the market and the putting into service of RBI systems not by law enforcement authorities, "as well as the use of other RBI systems, is not prohibited, but subject to the rules for high-risk AI systems". (European Commission, 2025)

The Act acknowledges the highly intrusive nature of RBI systems in public spaces by law enforcement to the rights and freedoms of individuals "to the extent that it may affect the private life of a large part of the population, evoke a feeling of constant surveillance and indirectly dissuade the exercise of the freedom of assembly and other fundamental rights", leading to biased and inaccurate results. (Recital 32, AIA) However in some cases, where the use is strictly necessary to achieve a substantial public interest, the importance of which outweighs the risks. (Recital 33, AIA)

According to Article 3(41) AI Act, a RBI system is "[a]n AI system for the purpose of identifying natural persons, without their active involvement, typically at a distance through the comparison of a person's biometric data with the biometric data contained in a reference database" , by

recognizing features such as "the face, eye movement, body shape, voice, prosody, gait, posture, heart rate, blood pressure, odour, keystroke characteristics" (Recital 15, AIA)

Biometric verification falls outside of the scope. (European Commission, 2025) Real time means that the system captures the biometric data instantaneously, or in any event without significant delay, requiring at least that the person has not left the place, (European Commission, 2025) and law enforcement purposes the prevention, investigation, detection and prosecution of criminal offences by law enforcement authorities or on their behalf. (European Commission, 2025) "Only when those other bodies or entities have been entrusted with a specific law enforcement task will their activities fall under the definition of 'law enforcement". (European Commission, 2025) It is worth highlighting that the GDPR already regulates the process of biometric data, as special categories of personal data, in Article 9. In addition, according to Recital 38, the rules of this Regulation "should apply as lex specialis in respect of the rules on the processing of biometric data contained in Article 10 of Directive (EU) 2016/680", the LED Directive, meaning that law enforcement authorities are confined to act only inside the limits of Article 5(1)(g) AIA.

As mentioned before, Article 5(1)(h) provides 3 exhaustive exceptions to the prohibition, balancing the protection of the right to privacy to the security of society (European Commission, 2025), under the condition that Member State legislation authorizes them. "Consequently, in the absence of Member State legislation authorising the use of real-time RBI for one or more of those objectives, such use is prohibited as from 2 February 2025". (European Commission, 2025) The use of those systems is permitted only where it is "strictly necessary to achieve a substantial public interest." (Recital 33, AIA)

The first exception (i) concerns targeted search for the victims of three serious crimes, abduction, trafficking, and sexual exploitation of human beings, and missing persons (whose search is made only for law enforcement purposes). The first scenario of the second (ii) exception is "the prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons", including critical infrastructure, for example when a "serious disruption and destruction of critical infrastructure (e.g., a power plant, water supply, or a hospital) may result in an imminent threat to the life or the physical safety". (European Commission, 2025) An imminent threat to life or physical safety is "a threat that can occur at any moment and requires

'immediate action to be taken'," while a substantial threat to physical safety "relates to serious bodily injuries". (European Commission, 2025) Ultimately, these terms are defined and assessed by each Member State, based on its national laws, in accordance with EU legislation. (European Commission, 2025) The second scenario is the prevention of a "genuine and present or genuine and foreseeable threat of a terrorist attack". The guidelines highlight that the terrorist threat level is determined at national level and varies from one Member State to another, but the notion of terrorism is "an autonomous notion of Union law and should therefore be assessed, in principle, independently of national definitions". (European Commission, 2025) Lastly, the third exception allows real time RBI for the localisation and identification of suspects or perpetrators of certain serious crimes referred to in Annex II AIA, such as terrorism, trafficking in human beings, sexual exploitation of children, and child pornography, rape, murder, grievous bodily injury, crimes within the jurisdiction of the International Criminal Court, environmental crime.

To prevent the abuse of the exceptions, which at some points may be vague, the Act in Articles 5(2) – (7) provides safeguards and conditions;

- The use of real time RBI "only to confirm the identity of the specifically targeted individual", taking into account the harm of not using the system, and the consequences to the fundamental rights of all persons concerned (Article 5(2), AIA),

- the national law authorizing the use must comply with comply with "necessary and proportionate safeguards and conditions […] in particular as regards the temporal, geographic and personal limitations" (Article 5(2), AIA),

- the use of real time RBI shall be authorized "only if the law enforcement authority has completed a fundamental rights impact assessment as provided for in Article 27 and has registered the system in the EU database according to Article 49" (Article 5(2), AIA) and "shall be subject to a prior authorization granted by a judicial authority or an independent administrative authority" (Article 5(3), AIA), except in situations of urgency. Compliance with the GDPR in carrying out data protection impact assessments is also necessary. "No decision that produces an adverse legal effect on a person may be taken based solely on the output of the 'real-time' remote biometric identification system" (Article 5(3), AIA),

- the use of such systems "shall be notified to the relevant market surveillance authority and the national data protection authority" (Article 5(4), AIA) which then "shall submit to the Commission annual reports on such use" (Article 5(6), AIA),

- Member States that decide to fully or partially authorize (or not authorize) such use "shall lay down in their national law the necessary detailed rules for the request, issuance and exercise of, as well as supervision and reporting relating to, the authorisations referred to in paragraph 3" (Article 5(5), AIA) and "notify those rules to the Commission" (Article 5(5), AIA).

It is true that the AIA makes an important step in prohibiting real time RBI, a very intrusive practice for the freedoms and fundamental rights of natural persons, to prevent forms of biometric mass surveillance. On the other hand, the significant risks of retrospective facial recognition are largely overlooked. They are considered high risk practices and are not banned by the AIA. (European Commission, 2025) It is claimed that other retrospective/post RBI are just as invasive as real time RBI and are used for mass surveillance. "Yet the AI Act makes a big error in claiming that the extra time for retrospective uses will mitigate possible harms" taking into account that the safeguards could easily be circumvented by law enforcement authorities. (Jakubowska et al., 2024) Civil societies that actively follow the AIA call for "a full ban on retrospective RBI by private and public actors" and "urge that the "significant delay" clause should be at a minimum of 24 hours after capture". (Access Now, 2024) Real time RBI uses for non-law enforcement purposes are not prohibited under the AIA but must in any case comply with data protection rules. (European Commission, 2025)

Remote biometric identification in publicly available places is regulated by the LED, since the purpose of the processing is law enforcement. Since the processing concerns biometric data, Article 8 for lawful processing and Article 10 for the prohibition of processing special categories of personal data apply. The controller must apply the principle of proportionality to determine the extent of necessary processing and also demonstrate the necessity of the processing for the performance of the task. It must be noted that it is difficult to justify such processing due to proportionality problems, considering the special categories of personal data involved, the large scale of processing, the strict exceptions of the prohibition, the application thereof must not circumvent the prohibition and the transparency problems.

In fact, the EDPB and the EDPS (2021) "call for a general ban on any use of AI for an automated recognition of human features in publicly accessible spaces- such as of faces but also of gait, fingerprints, DNA, voice, keystrokes and other biometric or behavioral signals- in any context" and in online spaces as well, for public and private use, and conclude that "taking into account the LED, the EUDPR and GDPR, the EDPS and EDPB cannot discern how this type of practice would be able to meet the necessity and proportionality requirements, and that ultimately derives from what are considered acceptable interferences of fundamental rights by the CJEU and ECtHR."

6.10 Exclusion from the scope of the AIA

It is important to note that Article 2 AIA provides for some exclusions from the scope, also relevant to the prohibited practices of Article 5 AIA; national security, defense and military purposes, judicial and law enforcement cooperation with third countries, research & development, personal non-professional activity, AI systems released under free and open source licenses. "This exemption introduces a significant loophole that will automatically exempt certain AI systems from scrutiny and limit the applicability of human rights safeguards envisioned in the AI Act". Even though the CJEU[66] interprets and confines the notion of 'national security', in practical terms, this could allow governments to justify the implementation of biometric mass surveillance systems under the pretest of national security, bypassing the safeguards outlined in the AI Act, avoiding the requirement for a fundamental rights impact assessment, and neglecting to ensure that the AI system meets high technical standards or does not discriminate against specific groups. The absence of clear national-level procedures to assess if the national security threat invoked by the government is in fact legitimate and serious enough to justify the use of the

[66] 7 Judgment of the Court of Justice of 6 October 2020, La Quadrature du Net and Others, C-511/18, C-512/18

system and whether the system is developed and deployed in accordance with fundamental rights, makes this exception very concerning. (Jakubowska et al., 2024)

6.11 Citations for chapter 6

European Commission (2024) *Annex to the Communication to the Commission: Approval of the content of the draft Communication from the Commission - Commission Guidelines on prohibited artificial intelligence practices established by Regulation (EU) 2024/1689 (AI Act)*. Brussels: European Commission

EDPB and EDPS (2021) *Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Brussels: European Data Protection Board and European Data Protection Supervisor.

EDPS (2021) *EDPS TechDispatch on Facial Emotion Recognition*. Brussels; EDPS

European Parliamentary Research Service (2020) *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*. Brussels: European Parliament.

Homo Digitalis (2024) *Πράξη για την τεχνητή νοημοσύνη – Ανάλυση διατάξεων για τις απαγορευμένες πρακτικές του άρθρου 5 του Κανονισμού 2024/1689*. Unknown: Homo Digitalis.

Access Now (2024). ***Upcoming Commission guidelines on the AI Act implementation***. Available at: https://www.accessnow.org/press-release/upcoming-commission-guidelines-on-the-ai-act-implementation/ [Accessed 30 March 2025].

Jakubowska E. et al (2024) *EU's AI Act fails to set gold standard for human rights.* Available at: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://edri.org/wp-content/uploads/2024/04/EUs-AI-Act-fails-to-set-gold-standard-for-human-rights.pdf [Accessed 30 March 2025].

7 AI ACT PRIVACY PROVISIONS IN RELATION TO DATA PROTECTION LAWS

7.1 Introduction

Apart from the prohibited practices, the AIA contains other provisions related to the protection of privacy and personal data. The most relevant section of the AIA regards high-risk AI systems. The AI systems that will not fall under the prohibited systems of Article 5 will most likely be characterized as high-risk systems under Article 6 in conjunction with Annexes I and III AIA. Such systems are subject to the strict requirements of Chapter 3 AIA, taking into account the purpose of the system (Article 8, AIA), which materialize the principles set in Recital 27 AIA. (privacy and data governance, transparency, non-discrimination etc.) Below I will examine the provisions of the Act that are most relevant to privacy and personal data protection, also in relation primarily to the GDPR and the LED.

7.2 Data governance and management

7.2.1 Data requirements

A significant provision of the AIA is the obligation regarding appropriate data governance and management (Article 10, AIA). As mentioned, the quality of the training data plays a vital role in the performance of AI systems. It is therefore crucial that high risk AI systems are trained with high quality data sets to ensure their safe performance and to avoid discriminative practices and follow the practices mentioned in Article 10(2) AIA. In particular, data sets used for training, validation and testing, including the labels, must be lawfully collected, relevant, suitable, sufficiently representative, unbiased and as error free as possible and complete considering the system's intended purpose. To ensure compliance with Union data protection law, such as the GDPR when personal data are processed, the systems must be transparent about the original purpose of the data collection, have the appropriate statistical properties, especially regarding the persons to whom the high-risk AI system is intended, with specific attention to the mitigation

of possible biases in the data sets. (Recital 67, AIA) (Article 10(1), 10(2), 10(3), AIA) They should also consider the characteristics that are particular to the specific geographical, contextual, behavioural or functional setting which the AI system is intended to be used. (Article 10(4), AIA) "To the extent the deployer exercises control over the input data, that deployer shall ensure that input data is relevant and sufficiently representative in view of the intended purpose of the high-risk AI system". (Article 27(4), AIA) The obligation to define the intended purpose of the personal data collected is connected to the purpose limitation principle of the GDPR (Article 5(1)(b)) and the data accuracy obligation of Article 5(1)(d) GDPR. (Data Protection Authority of Belgium, 2024)

7.2.2 Lawful processing

As previously stated, the processing of personal data in the context of development and deployment of AI systems must follow the data processing principles and be based on a legal basis according to Article 6 GDPR. The "legitimate interest" basis under Article 6(1)(f) GDPR usually causes the most confusion and for this reason the EDPB published an Opinion to clarify when controllers can rely on it as an appropriate legal basis. The controllers must use the three step test to assess whether they can base the processing of personal data for AI training under this legal basis, i.e. (1) Identifying the legitimate interest pursued by the controller or a third party, which must be lawful, clearly and precisely articulated, real and present. (2) "Analysing the necessity of the processing for the purposes of the legitimate interest(s) pursued", and "whether there is no less intrusive way of pursuing this interest" (also referred to as "necessity test"). (3) "Assessing that the legitimate interest(s) is (are) not overridden by the interests or fundamental rights and freedoms of the data subjects", taking into account the specific circumstances of each case and data subjects' reasonable expectations (also referred to as "balancing test")." (European Data Protection Board, 2024)

7.2.3 Algorithmic Discrimination

Providers of high-risk AI systems are explicitly exempted from the prohibition of Article 9 GDPR and may exceptionally process special categories of personal data to the extent that it is absolutely necessary for bias detection and correction purposes in accordance with Articles 10(2) and 2(7) AIA, subject to appropriate safeguards for the fundamental rights and freedoms of natural persons. The processing of sensitive data must be subject to the appropriate safeguards of the GDPR, the LED and the Data Protection Framework for EU Institutions. (Article 10(5), AIA) For example, in order to detect whether an algorithm used in the employment sector discriminates based on ethnicity, the organization in question would, in principle, need to know the ethnicity of the job applicants. Since in circumstances when the AIA and the GDPR collide the latter prevails, an appropriate legal basis for such processing according to the GDPR must be established. The "substantial public interest ground'[67] of 9(2)(g) GDPR is considered an appropriate basis for this processing. "In this context, the AI Act would serve as the relevant Union law, and fighting discrimination would be the relevant public interest." (European Parliamentary Research Service, 2025) The personal data processing principles and other GDPR provisions, such as security and necessity requirements and the grounds of Article 9 GDPR must be applied accordingly in any case.

However, it is noted that discrimination could also arise from data that are not considered special category personal data but merely personal data (e.g. gender or age) which shall be processed under the broader legal grounds of Article 6 GDPR (e.g. legitimate interest) "In general terms, shared uncertainty appears to prevail as to how the AI Act's provision on the processing of special categories of personal data for avoiding discrimination should be interpreted. The GDPR, which imposes limits on the processing of special categories of personal data, might prove restrictive in a context dominated by the use of AI in many sectors of the economy, and faced with the mass processing of personal and non-personal data." (European Parliamentary Research Service, 2025)

[67] Article 9(2)(g) GDPR permits processing of special categories of data when "necessary for reasons of substantial public interest, on the basis of Union or Member State law which shall be proportionate to the aim pursued, respect the essence of the right to data protection and provide for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject"

7.3 Automated decisions - Human Oversight

Another important obligation of AIA for high-risk AI systems is human oversight. The so called "human in the loop effect" is necessary to ensure that fundamental rights are protected "when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse" (Article 14(2), AIA), to guarantee legal and ethical compliance and system transparency. Article 14 AIA requires providers to ensure that high risk AI systems are "designed and developed in such a way, including with appropriate human-machine interface tools, that they can be effectively overseen by natural persons during the period in which they are in use." (Article 14(1), AIA) Plus, Article 26(1) AIA requires deployers of such systems to "take appropriate technical and organisational measures to ensure they use such systems in accordance with the instructions for use accompanying the systems", including with respect to human oversight. The systems must have in-built operation constraints, be designed and developed in such a way so that natural persons can duly monitor their operations and include mechanisms to guide the person assigned to oversee it, while those persons must have the necessary competence and training. In case of remote biometric identification no action or decision is taken by the deployer unless the output is confirmed by 2 competent natural persons, except for "systems used for the purposes of law enforcement, migration, border control or asylum, where Union or national law considers the application of this requirement to be disproportionate". (Article 14(5), AIA) Affected persons subject to a decision which is taken by the deployer on the basis of the output from a high-risk AI system listed in Annex III, "shall have the right to obtain from the deployer clear and meaningful explanations of the role of the AI system in the decision-making procedure and the main elements of the decision taken". (Recital 86, AIA)

On the other hand, Article 22(1) GDPR provides that as a rule, there is a general prohibition on fully automated individual decision-making and profiling that produces legal effects concerning the data subject and provides them with the right to contest the decision. The prohibition covers

decisions where there is no human involvement meaning that "the controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision. As part of the analysis, they should consider all the relevant data." (Article 29 Data Protection Working Party, 2018) There are exceptions where ADM is permitted[68] (Article 22(2), GDPR) provided that the controller implements "suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision." (Article 22(3), GDPR). In any case, under the GDPR data subjects enjoy the rights of Chapter III GDPR including the right to be informed of Articles 13(2) (f) and 14(2) (g) and the right of access Article 15(1)(h). Therefore, a controller making automated decisions as described in Article 22(1) GDPR must "tell the data subject that they are engaging in this type of activity, provide meaningful information about the logic involved and explain the significance and envisaged consequences of the processing." (Article 29 Data Protection Working Party, 2018) Overall under the GDPR individuals enjoy more rights, including the right to contest the decision.

When comparing the provisions of the two Regulations we are led to the conclusion that they don't fully align. According to the AIA there must always be a competent natural person to oversee the system and "decide not to use the AI system or to override or reverse the output" (Article 14(4), AIA) (Sarra, 2025) It appears that the AIA sets a higher threshold to what is considered human oversight in comparison with the GDPR, since it includes a wider range of activities. Therefore, there will be cases where a decision will be considered fully automated according to the AIA but not according to the GDPR, rendering Article 22 GDPR inapplicable. In this scenario, the individual is entitled to a right to explanation of Article 86 AIA as mentioned above but will not have the rights granted by Article 22 GDPR, including the right to contest the

[68] "if the decision:
(a) is necessary for entering into, or performance of, a contract between the data subject and a data controller;
(b) is authorised by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or
(c) is based on the data subject's explicit consent."

decision. Sarra argues that a new interpretation of Article 22 GDPR is required. (2025) Also it should be added that under the AIA the requirement for human oversight regards high-risk systems, thus making it possible "for an AI system classed as low or minimal risk to make a solely automated decision". (EU AI Act, 2025) It is worth noting that the existence of a right to explanation of automated decision making pursuant to Article 22 GDPR has been confirmed by the CJEU[69].

Lastly, the LED is more flexible providing that automated decision making, including profiling, is prohibited when it causes adverse legal effects concerning the data subject or significantly affects him or her, "unless authorised by Union or Member State law […] and which provides appropriate safeguards for the rights and freedoms of the data subject, at least the right to obtain human intervention on the part of the controller", without the possibility for the data subject to challenge the decision. (Article 11(1), LED) Nevertheless, paragraph (3) states that profiling that results in discrimination against natural persons on the basis of special categories of personal data shall be prohibited.

7.4 Security Measures

The AIA further provides robust security measures for high-risk AI systems, demanding that they must be designed to be accurate, robust and secure and perform consistently throughout their lifecycle. (Article 15(1), AIA) They should be resilient to errors and faults, have backup plans in place (Article 15(4), AIA) and "be resilient against attempts by unauthorised third parties to alter their use, outputs or performance by exploiting system vulnerabilities" meeting an appropriate lever of cybersecurity (Article 15(5), AIA) Also, "High-risk AI systems that continue to learn after being placed on the market or put into service shall be developed in such a way as to eliminate or reduce as far as possible the risk of possibly biased outputs influencing input for future operations (feedback loops), and as to ensure that any such feedback loops are duly

[69] CJEU 27 February 2025, no. C-203/22, Dun & Bradstreet Austria

addressed with appropriate mitigation measures." (Article 15(1), AIA) (Article 15(4), AIA) To address these risks, the AIA requires adequate risk management throughout the AI's lifecycle according to Article 9 AIA. Deployers must identify and analyze potential risks to health, safety, or fundamental rights, estimate and evaluate these risks, and adopt measures to manage them, in combination with human oversight, data governance and transparency obligations. Equivalent requirements are provided for GPAI in Article 55(1)(d) AIA. Nolte et al. argue that clearer specifications are needed through technical standards and EU Commission guidelines to address the legal challenges and shortcomings. They shall focus on "i) identify the technical requirements associated with vague legal terms; ii) defining the required level of 'robustness' and 'cybersecurity' and other concepts such as 'consistency'; iii) defining the requirements for evaluating and assessing AI systems and its components; and iv) pay attention to some aspects that are not explicitly regulated, such as feedback loops in offline systems in Art. 15(4) AIA or organizational measures to ensure 'cybersecurity' in Art. 15(5) AIA". (2025)

The Act builds upon the obligation of secure processing of the GDPR (Article 32, GDPR). The latter requires controllers and processors of personal data to implement appropriate technical and organizational measures (TOMs) like data encryption, access controls penetration testing and logging and auditing. Following that, the AIA recognizes the specific risks of AI and sets additional security "measures tailored to the specific vulnerabilities of the AI system, such as data validation and quality assurance", emphasizing the need for quality and integrity of the data used to train and operate the AI system. (Data Protection Authority of Belgium, 2024) The same obligations are provided by Article 19 LED.

7.5 Conformity assessment and Fundamental Rights Impact Assessment

The AIA imposes a conformity assessment obligation, to ensure accountability by the provider for developing high-risk AI systems, before the system is placed on the market or made available or substantial modifications take place. According to Article 43 AIA the provider must

demonstrate compliance with the requirements of Title III, Chapter 2 AIA; Risk management system, data governance, technical documentation, record keeping, transparency and provision of information, human oversight, accuracy, robustness, and cybersecurity. "Conformity assessments are not risk assessments but rather demonstrative tools that show compliance with the EU AI Act's requirements." (Clark et al., 2024)

Prior to deploying a high-risk AI system deployers that a) are bodies governed by public law, b) are private entities providing public services, c) of "AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score, with the exception of AI systems used for the purpose of detecting financial fraud" and d) "AI systems intended to be used for risk assessment and pricing in relation to natural persons in the case of life and health insurance" (Points 5 (b) and (c) of Annex III) must perform an assessment of the impact on fundamental rights (FRIA) that the use of such system may produce. (Article 27, AIA) "Whilst risks related to AI systems can result from the way such systems are designed, risks can as well stem from how such AI systems are used. Deployers of high-risk AI systems therefore play a critical role in ensuring that fundamental rights are protected, complementing the obligations of the provider when developing the AI system." (Recital 93, AIA)

The FRIA shall describe how and when the system will be used, who it might affect, and what risks it might pose. Taking into account that fundamental rights cover a broad range and there is extended CJEU case law thereof, it requires an in-depth knowledge to assess the impact of high-risk systems on these rights. It remains to be seen how the AI Office will address these challenges in its upcoming FRIA Template Questionnaire. (Article 27(5), AIA).

Nonetheless, in regard to the fundamental right to privacy, deployers of high-risk AI systems shall comply with their obligation to carry out a data protection impact assessment (DPIA) under Article 35 GDPR or Article 27 LED. (Article 27(9), AIA) "If any of the obligations laid down in this Article is already met through the data protection impact assessment conducted pursuant to Article 35 of Regulation (EU) 2016/679 or Article 27 of Directive (EU) 2016/680, the fundamental rights impact assessment referred to in paragraph 1 of this Article shall complement that data protection impact assessment." (Article 27(4), AIA) A DPIA is mandatory whenever processing of personal data is likely to result in a high risk to the rights and freedoms of individuals. It is

required particularly in cases of "systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling", processing of sensitive data on a large scale and systematic monitoring of public areas on a large scale. (Article 35(3), GDPR)

7.6 Transparency Obligations for Providers and Deployers of Certain AI Systems

Article 50 AIA imposes additional transparency obligations to providers and deployers of certain AI systems, high risk and limited risk, in order to make sure that individuals are aware that they are engaging with an AI system to "prevent misinformation and manipulation at scale, fraud, impersonation and consumer deception" and to make the traceability of information possible (Recital 133, AIA). More specifically, providers must ensure that AI systems intended for direct interaction with natural persons are designed and developed to clearly inform those individuals that they are engaging with an AI, unless it is clearly obvious or it is for law enforcement purposes. (Article 50(1), AIA) AI systems, including GPAI systems, that produce synthetic audio, image, video or text content or deepfakes must mark their outputs as artificially generated and (Article 50(2), AIA) (Article 50(4), AIA) and must inform natural persons when they use AI for emotion recognition or biometric categorization and shall process the personal data in accordance with the GDPR, LED and EUDPR, as applicable, unless it's for law enforcement purposes (Article 50(3), AIA). For the practical implementation of this provision the Commission is expected to develop guidelines according to Article 96 AIA. The information mentioned above must be communicated to the natural persons concerned in a clear and easily recognizable manner no later than at the time of their first interaction or exposure to the AI system. (Article 50(5), AIA)

7.7 AI Sandboxes

AIA introduced a tool to promote innovation while providing the safeguards needed to build trust and resilience, the AI regulatory sandboxes, allowing businesses to experiment with new

and innovative products, services or businesses (Article 57, AIA) under the supervision of national data protection authorities and other national or competent authorities where personal data are being processed (Article 57(10), AIA). Article 59 is dedicated to the further processing of personal data for developing certain AI systems for the public interest, such as public safety, health, environmental protection, energy sustainability, transport safety, and public administration. As a derogation from the GDPR (Article 2(7), AIA), "In the AI regulatory sandbox, personal data lawfully collected for other purposes may be processed solely for the purpose of developing, training and testing certain AI systems in the sandbox", while ensuring appropriate safeguards. The EDPB and EDPS note that the purpose limitation principle of the GDPR must be respected to avoid any inconsistences and conflicts. (EDPB and EDPS, 2021)

7.8 Rights granted

The AI Act, with a couple of exceptions, does not confer any rights to individuals who are adversely affected by non-compliance with the Act to seek remedies for breaches, since it is considered primarily a product safety law. One exception is lodging a complaint with a Market Surveillance Authority for an infringement of the Act without a provision to award remedies (Article 85, AIA). The second exception gives individuals the "right to obtain from the deployer clear and meaningful explanations of the role of the AI system in the decision-making procedure and the main elements of the decision taken" if a decision taken by the deployer on the basis of the output from a high-risk AI system has an adverse impact on their health, safety or fundamental rights. (Article 86(1), AIA) This provision applies only to the extent that this right is not otherwise provided for under Union law. (Article 86(3), AIA) The right to explanation for automated decisions by AI systems not characterized as high risk is not obligatory.

On the other hand, the GDPR is a fundamental rights law that grants rights to individuals in relation to the processing of their personal data. (Clark et al., 2024) Under the GDPR data subjects have the rights mentioned in Section 2 and more specifically the right to be informed (Articles 12-14, GDPR), right to access (Article 15, GDPR) right to rectification (Article 16, GDPR) right to

erasure (Article 16, GDPR) the right to restrict processing (Article 18, GDPR) the right to data portability (Article 20, GDPR), the right to object (Article 21, GDPR) and right regarding automated individual decision making (Article 22, GDPR). Additionally, they have the right to lodge a complaint with a supervisory authority (Article 77 GDPR) and seek effective judicial remedy (Article 78 GDPR).

As already mentioned, the CJEU in case C-203/22, confirms that the data subject has the right to explanation of algorithmic decisions in a concise, transparent, intelligible and easily accessible manner as part of the right of access (Article 15(h), GDPR and Article 22 GDPR). This includes the provision of meaningful information about the logic involved, covering "all relevant information concerning the procedure and principles relating to the use, by automated means, of personal data with a view to obtaining a specific result." The complexity of the system cannot be used as an excuse for "relieving the controller of the duty to provide an explanation". It also underlines that "the main purpose of the data subject's right to obtain the information provided for in Article 15(1)(h) of the GDPR is to enable him or her effectively to exercise the rights conferred on him or her by Article 22(3) of that Regulation, namely the right to express his or her point of view on that decision and to contest it." This judgement is of great significance establishing for the first time algorithmic transparency in the GDPR. (Rossetti, 2025)

7.9 MSAs and Fines

Member States shall appoint Market Surveillance Authorities at national level for the purpose of supervising the application and implementation of the AI Act. The EDPB, amongst others, proposes DPAs for the role since they already have experience and expertise in addressing the impact of AI on fundamental rights, in particular the right to protection of personal data. EDPB Deputy Chair Irene Loizidou Nicolaidou said: "DPAs should play a prominent role in enforcing the AI Act as most AI systems involve processing of personal data. I strongly believe that DPAs are suitable for this role because of their full independence and deep understanding of the risks of AI for fundamental rights, based on their existing experience." (EDPB, 2024)

DPAs as MSAs will have the power to request corrective measures and impose high fines for violations of the Act. (Article 99, AIA). Under the GDPR, DPAs as Supervising Authorities have similar powers. (Article 58, GDPR). Nevertheless, a recent analysis of EDPB statistics conducted by noyb revealed the low activity of most national DPAs. "The data shows that, on average, merely 1.3% of cases before DPAs result in a fine." Noyb insists that these numbers are concerning since surveys show that it is precisely monetary fines that motivate companies to comply with the law. (noyb, 2024) This is an important observation that must be taken into account when discussing the enforcement of the GDPR and the AI Act.

The EDPB in a recent Opinion addressed "the consequences of the unlawful processing of personal data in the development phase of an AI model on the subsequent processing or operation of the AI model", providing some considerations, while pointing out the discretionary powers of the SAs. It sets the approach on how SAs should treat cases of unlawful processing depending on whether the data is retained in the model and is processed/deployed by the same controller or a different one or the data is anonymized. The EDPB points out that the demonstration of accountability on behalf of the controller plays a key role in the assessment. Lastly, the EDPB reiterates that "the measures taken by SAs under the GDPR are without prejudice to those taken by competent authorities under the AI Act". (European Data Protection Board, 2024)

7.10 Citations for chapter 7

European Parliament and Council Regulation (EU) 2024/1689 of the 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) (Text with EEA relevance) Available at; https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free

movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) Available at; https://eur-lex.europa.eu/eli/reg/2016/679/oj/eng

Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA Available at; https://eur-lex.europa.eu/eli/dir/2016/680/oj/eng

European Parliamentary Research Service (2025) *Algorithmic discrimination under the AI Act and the GDPR.* Brussels; European Parliamentary Research Service.

European Parliamentary Research Service (2020) *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*. Brussels: European Parliament.

EDPB and EDPS (2021) *Joint Opinion 5/2021 on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*. Brussels: European Data Protection Board and European Data Protection Supervisor.

Data Protection Authority of Belgium (2024) *Artificial Intelligence Systems and the GDPR: A Data Protection Perspective.* Brussels: Data Protection Authority of Belgium.

Article 29 Data Protection Working Party (2018) *Guidelines on Automated Individual Decision-making and Profiling for the Purposes of Regulation 2016/679*. Brussels: European Commission.

European Data Protection Board (2024) *Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models.* Brussels: European Data Protection Board.

Sarra, C. (2025) 'Artificial Intelligence in Decision-making: A Test of Consistency between the "EU AI Act" and the "General Data Protection Regulation"', *Athens Journal of Law*, 11(1), pp. 45-62. Available at: https://www.athensjournals.gr/law/2025-11-1-3-Sarra.pdf [Accessed: 7 April 2025].

Nolte, H., Rateike, M. and Finck, M. (2025). 'Robustness and Cybersecurity in the EU Artificial Intelligence Act'. *Unpublished.* Available at: *arXiv:2502.16184v1 [cs.AI]* [Accessed: 17 April 2025].

Clark, J., Demircan, M. and Kettas, K. (2024) '*Europe: The EU AI Act's relationship with data protection law: key takeaways. DLA Piper*, 25 April. Available at: https://privacymatters.dlapiper.com/2024/04/europe-the-eu-ai-acts-relationship-with-data-protection-law-key-takeaways/ [Accessed: 11 April 2025].

EU AI Act (2025). **Key Issues, Interplay with GDPR**. Available at: https://www.euaiact.com/key-issue/6 [Accessed 11 April 2025].

EDPB (2024). **EDPB adopts statement on DPAs' role in the AI Act framework and the EU-U.S. Data Privacy Framework FAQs.** Available at: https://www.edpb.europa.eu/news/news/2024/edpb-adopts-statement-dpas-role-ai-act-framework-eu-us-data-privacy-framework-faq_en [Accessed 11 April 2025].

noyb (2024). **Data Protection Day: Only 13% of cases by EU DPAs result in a fine**. Available at: https://noyb.eu/en/data-protection-day-only-13-cases-eu-dpas-result-fine [Accessed 15 April 2025].

Rossetti, S. (2025) 'The Court of Justice of the European Union confirms the existence of the right to explanation of automated decision-making', European Law Blog, 7 April. Available at: https://www.europeanlawblog.eu/pub/lwchuopd/release/1 [Accessed 23 April 2025].

8 CONCLUSION

8.1 Shortcomings

AI is reshaping societies, economies and individual lives at a rapid pace. While it brings considerable benefits such as improved efficiency, personalization and innovation, it also introduces significant challenges particularly in relation to privacy and the protection of personal data. This thesis has explored these concerns with a specific focus on the implications of AI systems and their regulatory responses in the European level.

The objective of this paper was to provide a critical assessment of whether the AIA, the EU's flagship AI regulation, is both relevant and effective in safeguarding the fundamental right to privacy in the era of AI technologies. The AIA represents a critical step towards creating a harmonized framework to ensure that AI is trustworthy, ethical and aligned with fundamental rights. However, despite its broad scope, the Act does not fully address the privacy and data protection risks associated with AI technologies. While the AIA introduces prohibitions and requirements for AI systems, many privacy threats remain insufficiently covered.

First of all, the definition of an AI system (and model) by the Act has been heavily criticized as vague and overly broad making it difficult to distinguish an AI system from traditional automation software. This lack of precision undermines legal clarity for the operators of AI systems and for the enforcement authorities.

One of the most significant shortcomings of the AIA is the narrow scope of its prohibited practices. While the prohibitions of Article 5, such as real time biometric identification in public spaces for law enforcement, social scoring by public authorities and manipulative techniques that impair user autonomy, are important, as explained in chapter 6, they apply only in very specific situations. Many harmful uses of AI will likely not meet the threshold and be left out of the scope of the prohibition. For example, emotion recognition and biometric categorization are only prohibited in schools and workplaces even though they raise serious privacy concerns in other settings, and other risky practices like profiling, mass surveillance or predictive policing may still be allowed as long as they meet certain technical rules. This narrow focus means that the Act does

not fully address how AI can harm privacy when used in more subtle or widespread ways. In the end, the list of prohibited practices is too limited to offer strong protection against the broad range of privacy risks that AI can create.

Plus, the AIA tends to approach these risks from a technical and systemic perspective focusing on conformity assessments and product safety obligations rather than granting individual rights (except from the right to explanation) or providing enforceable mechanisms of redress.

Another concern that is raised about the scope of the Act regards high risk AI systems. Given that Article 6 AIA will be entered into force on 2 August 2027[70], existing high risk systems and systems developed in the meantime will not be regulated. This is an important gap that the AIA fails to take intro consideration.

Lastly, the implementation and supervision of the enforcement of the Act has been assigned to a multitude of European and national authorities each with different responsibilities. The cooperation of the European Artificial Intelligence Board, the AI Office, the Scientific Panel and the Advisory Forum, with the national Market surveillance authorities and Notifying authorities, may prove ineffective and dysfunctional.

Overall, existing legal frameworks, notably the GDPR and the LED, already provide a more robust foundation for privacy protection. While they don't specifically regulate AI, they are specifically designed to regulate the processing of personal data which are the core of most AI systems. Article 22 GPDR, on automated decision making, is directly relevant to many of the most pressing risks associated with AI. The Act, by comparison, does not offer equivalent rights or safeguards nor does it directly regulate how AI systems process personal data or introduce meaningful constraints on profiling, inference or behavior manipulation, unless such system fall under the narrow scope of the prohibited practices. The GDPR and LED include key principles like data minimization, transparency and lawfulness which are directly applicable to AI systems that process personal data. Many of the obligations that the AIA poses to operators, like implementing accountability measures, security measures and impact assessments are already covered.

[70] Article 113 AIA

8.2 The complimentary nature of the AIA

This analysis demonstrates that while the Act reinforces existing privacy norms and introduces important provisions, it cannot replace the GDPR or LED in regulating personal data processing and in securing individual's rights. Instead, the Act should be seen as complementary, offering a structural approach and governance, with a focus on product safety, while relying on established data protection regimes to safeguard individual rights. The EDPB highlights that since EU data protection law is fully applicable to the processing of personal data involved in the lifecycle of the AI systems, the AIA and data protection legislation need to be, in principle, considered as complementary and mutually reinforcing instruments. It also stresses the importance of the interaction between the MSAs with DPAs and other authorities protecting fundamental rights. (2024)

In light of this, future regulatory efforts must strengthen the interplay between AI specific legislation and general data protection laws. This includes enhancing transparency, accountability and oversight mechanisms particularly for high risk and general purpose AI systems.

Ultimately the effective regulation of AI and protection of privacy will not depend only on legal frameworks but also on political will, public awareness and technological solutions that embed privacy by design and by default into AI systems.

8.3 Citations for chapter 8

European Data Protection Board (2024) *Statement 3/2024 on data protection authorities' role in the Artificial Intelligence Act framework.* Brussels: European Data Protection Board.

European Parliament and Council Regulation (EU) 2024/1689 of the 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) (Text with EEA relevance) Available at; https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng.