# From Regulating **Human Behaviour** to Regulating **Data**

Editors
**B. van der Sloot, G. Monti & F. Bostoen**

OPEN PRESS TiU

FROM REGULATING HUMAN BEHAVIOUR TO REGULATING DATA

# FROM REGULATING HUMAN BEHAVIOUR TO REGULATING DATA

EDITORS **B. VAN DER SLOOT | G. MONTI | F. BOSTOEN**

OPEN PRESS TiU

Open Press Tilburg University
Tilburg, The Netherlands



Design, typesetting, copyediting: LINE UP boek en media bv

ISBN: 9789403773148
DOI: https://doi.org/10.26116/k295-r264

# Table of contents

PART

# I

# INTRODUCTION

CHAPTER

# I

# Introduction

**Bart van der Sloot,**

**Giorgio Monti &**

**Friso Bostoen**[1]

https://doi.org/10.26116/qh2q-bc70

1 Associate, Full and Assistant Professor, Tilburg Institute of Law, Technology, and the Society (TILT), Tilburg Law School (TLS), Tilburg University.

## 1. Introduction

This chapter lays the groundwork for this book, which contains research on the intersection between innovation, regulation and democracy. Section 2 describes the background of the research agenda developed by the Tilburg Institute for Law, Technology and Society (TILT), which ultimately resulted in this edited volume. Section 3 outlines the research questions that ground this research agenda, as well as this book. Section 4 homes in on nine specific projects developed under the umbrella of this research agenda and discusses the most important perspectives, methodologies and questions that drive these projects. Finally, section 5 provides an overview of the content of this book. It contains answers, solutions and proposals but also – perhaps more often – further complications and dilemmas these projects have exposed.

## 2. Background of this Book

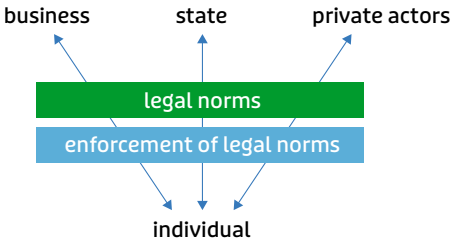


**FIGURE 1.** *The classic, human-centric regulatory paradigm*

Historically, regulators have targeted human behaviour, aiming to achieve public policy goals by influencing how both natural and legal persons act and how their activities affect others and society.[2] Law regulates human behaviour in various relationships. Classically, public law regulates relationships between citizens and states, while private law regulates relationships between consumers and industry as well as relationships between individuals and between businesses.

Under the influence of, inter alia, globalisation, privatisation and digitisation, changes to this classic approach have taken place over the past decades. Polycentric governance, shifting power relations, and complex networks of regulatory structures have forced regulators to broaden their toolbox and include instruments such as soft-law, self-reg-

2 Sections 2–4 are based on TILT's Sectorplan description, which in turn was based on text contributions from various Sectorplan members and composed by, among others, Ronald Leenes and Bert-Jaap Koops.

ulation and certification. What has been left unaltered, however, is the focus on human behaviour and relationships. In the 21st century's data-driven society, with the rise of artificial intelligence (AI), big data and robotics, the classic regulatory paradigm has been challenged in fundamental ways:

- The extent to which public policy goals can adequately be achieved by regulating human behaviour in the context of human relationships is increasingly uncertain. In contemporary debates on regulatory challenges in the context of data science and algorithmic decision-making, the focus is on how data behaves and how technological systems operate, rather than seeking to understand individual or collective human behaviour.
- Increasingly, rules are automatically enforced through data systems that can self-adapt based on feedback loops. The clear distinction between norm-setting and norm enforcement (i.e. substantive law versus procedural law; the content of legal rules versus the enforcement of legal rules through the judicial system) is thereby blurred. Human discretion and human interpretation of rules as key features of the legal system may over time fade to the background.

These two interrelated developments imply that the regulatory paradigm is gradually shifting from a human-centric paradigm to a data-centric paradigm, raising fundamental questions on the formulation and enforcement of norms and how to shape adequate checks and balances in regulatory processes.

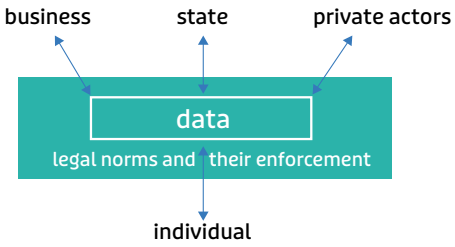


**FIGURE 2.** *The new, data-centric regulatory paradigm*

## 3. Research Agenda

This is why TILT started a research programme, funded by the Dutch government, with this guiding question: how can the shift from a human-centric regulatory paradigm to a data-centric regulatory paradigm be mapped, understood and, if possible, shaped?

Three research lines were developed to answer this question, which are based on three key elements of regulation.

- Stream 1 has as its overarching research question: how can **rules** be formulated and enforced when regulation shifts from a focus on human behaviour and human relations to data relations and the behaviour of data systems?
  - How does data-driven decision-making affect the legal system?
  - What legal and social challenges arise as people interact with data and AI systems within the legal system?
  - What rules are appropriate to regulate digital competitive advantage in a data-driven platform economy?
- Stream 2 has as its overarching research question: what **concepts** can be used in regulation, when regulation shifts from a focus on human behaviour and human relations to data relations and the behaviour of data systems?
  - What concepts are suitable for regulating data relations, when all data are personal data and the General Data Protection Regulation (GDPR) seems to have become the law that governs everything?
  - What concepts are suitable for capturing privacy protection in regulation to adequately protect trust and identity in the context of data-driven technologies?
  - How can a common understanding of the concept of consent in data protection and in contract law guide the legitimate processing of personal data and the entering into a contract?
- Stream 3 has as its overarching research question: what **values** are vital when regulation shifts from a focus on human behaviour and human relations to data relations and the behaviour of data systems?
  - What is the role of human autonomy in an era of machine autonomy and data-driven decision-making?
  - How can economic regulation safeguard the autonomy of individuals and businesses vis-à-vis digital giants in a data-driven society?
  - What regulatory, social and ethical challenges emerge when healthcare personnel interact with data and AI systems within the healthcare setting, and how can code contribute to supporting a normative framework?

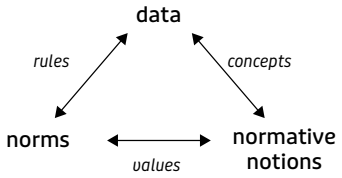


**FIGURE 3.** *The data-oriented regulatory triangle*

## 4. The Projects

To answer these questions, TILT started nine interdisciplinary, cross-sectional research projects, each with a dedicated team of researchers with backgrounds in law, philosophy, sociology, technology, psychology and economics.

### Project 1 – The Impact of Data and AI Systems within the Legal System

The first project focusses on the question of how data-driven AI and decision-making affect the legal system and aims to develop a fundamental conceptual framework for comparing the classic, human-centred legal system to a new, data and AI-centred legal system as well as to apply legal data science techniques for measuring (qualitatively and quantitatively) the effects of (big) data and AI on the legal system. In order to understand the impact of AI systems, the project team develops and evaluates methodologies for assessing the effects such systems have on the legal system. This has an epistemological component (where should we look to understand the effects of [big] data and AI on the legal system and on legal practice and are we asking the right questions about these effects?) as well as a methodological component (how can we successfully employ use of legal data science techniques [e.g. natural language processing] on legal research questions to assess the impact of AI on the legal system?).

Three case studies, which focus on the use of AI for immigration decision-making by immigration services, the use of AI for decision-making on environmental permissions, traffic fines and legal oversight of the vulnerable, and the use of AI by law enforcement agencies, yield the concepts and questions that underpin our understanding of possible effects of AI on the legal system. For example, the behaviour of AI systems can be evaluated from both a technical and a legal perspective. In AI and machine learning, there are technical evaluation metrics for algorithms – such as accuracy, precision and computational complexity. Comparatively, from a legal perspective, one evaluates the impact on the rights of the subject and due process. AI and data-driven decisions require rethinking the conceptual framework. For instance, does the fact that an algorithm can predict a human judge's decisions with a high accuracy mean that we can, in the future, give the algorithm the power to make similar decisions?

### Project 2 – Integrating AI into the Human-Centric Legal System

The driving question for this project is what legal and social challenges are arising as people interact with data and AI systems within the legal system. This project focusses on understanding problems of human interaction with AI in the legal system and interrogates how notions of fairness and justice are operationalised within the legal sector.

This project builds on Project 1 and looks at whether automation bias (a tendency not to question machines' guidance) occurs, what level of discretion is left to decision-makers and whether this is changing with the embedding of AI, how the quality and risks of decision-making processes based on AI are evaluated, and how explainability is conceptualised in different applications and environments. The project team employs methodologies and insights from media studies, which is currently one of the most important 'home disciplines' for studies of AI and fairness, and which increasingly produces researchers who can work on theoretical questions incorporating technical understandings of AI systems.

### Project 3 – Competitive Advantage in the Digital Platform Economy

Society being data-driven increases the risk of interferences with fundamental rights. Fundamental rights infringements by tech giants are more likely to be consented to or ignored by citizens, since they have no market alternative to the services these entities provide. The absence of competition thus renders existing legislation, such as the GDPR and the Charter of Fundamental Rights (CFR), insufficiently capable of guaranteeing fundamental rights in a digital society. Competition law's adaptability to economic reality allows it to better tackle the digital transformation because it does not depend solely on a centralised authority, being fully open to enforcement by private parties. It has the tools to consider the exchange of personal data against services on digital platforms as a market transaction, and it provides a framework to balance innovation, regulation and ethical aspects of new technologies.

The third project focusses on two strands of research. On the one side, there are questions of market power related to the lack of short-term alternatives to digital services: the tipping of markets with network effects, the data necessary to enter those markets, the reduction of choice and data protection, the exclusion of competitors using the same platform. On the other side, there are questions of strategic action across the digital sector: the use of shared assets like platforms and intellectual property, the profiling of consumer preferences through personal data, the mobilisation of data and other resources for future innovation, and the influence over regulatory and infrastructure design.

### Project 4 – Understanding Data as an Economic Good and Data Governance Through Property

The fourth project – situated on the intersection of law, economics and political economy – explores the concept of data as an economic good as a possible alternative analytical framework for informing regulation of data relationships. Conventional ideas about information as a public good are transferred by some to the modern context of data without

question. The recent leaps in information technology, however, challenge these ideas, as technology fundamentally alters the nature of something as a good, for example, by making it excludable or subtractable. This project explores property rights to data. For instance, market mechanisms enabled via private property are traditionally considered a preferred solution for governing a private but not a public good. Collective property rights may provide a legal tool to limit access to data, which is essential for sustainable data use. This project aims to understand the nature of data as an economic good within a data-driven society through the lens of data relationships, as well as to formulate strategies for regulating data relationships, informed by the understanding of data as a good.

### Project 5 – Exploring the Interaction Between Privacy, Trust and Identity in the Data-Driven Age

Under the current paradigm, privacy is predominantly seen as an individual right, protecting individual interests like personal autonomy. The GDPR aims to empower data subjects by providing them with more control over their data, among other things. However, there is a growing gap between the legislative approach and the everyday privacy experience of data subjects. First, data subjects do not experience control over their data. Their privacy expectations are predominantly based on the trust they have in the apps and devices they use. Second, data subjects are increasingly confronted with unwanted information about themselves, illustrating the ineffectiveness of the current legal regime to achieve one of its major goals, namely the protection of digital identities. That is why this project on the interrelationship between privacy, trust and identity focusses on two points in particular:

- First, data subjects seldom make active decisions on the sharing of personal data, and if they do, this process is characterised by confusion, dependency and vulnerability rather than by autonomy, confidence and control. Moreover, technological applications are increasingly designed to invoke trust, regardless of whether that trust is justified or not. While user perceptions about privacy and the limits of exercising meaningful control in data-driven environments have been studied empirically, theory-building on such experiential aspects of privacy is mostly lacking.
- Second, in the data-driven environment, the individual will be confronted with unwanted information about themselves. Although the current legal paradigm does grant the individual a right to withhold access to personal data from others, it is silent on the question of how the individual should be protected vis-à-vis information communication to herself. Consent and control cannot be the only mechanism used to solve this dilemma.

## Project 6 – Consent and Contracts

An increasing number of contracts in the data economy requires that personal data is provided in some way by the buyer of goods and services. Consumers agree to provide the suppliers of those goods and services with all kinds of data about themselves and their preferences. However, the use of data often goes beyond what consumers thought they had agreed to and not always in beneficial ways. Data protection legislation provides for principles relating to processing of personal data, such as lawfulness, fairness, purpose limitation and data minimisation and for a set of rules that specify rights and obligations relating to data processing. Consumers often provide their consent and agree to uses of data that are undesirable or that they would not have wanted to agree to if they had been entirely free to choose. However, contracts can mostly only be concluded by agreeing with the general terms and conditions of the supplier. These terms and conditions can generally not be partially accepted.

One particularly striking problem is the inclusion of terms that give the supplier the right to unilaterally change the terms and conditions. It is further unclear whether the agreement expressed by the consumers qualifies as valid consent under the data protection regulatory framework. In which cases can a data controller rely on the ground that legitimates data processing when it is necessary for the performance of a contract with the data subject? This ground should be interpreted narrowly, covering situations when processing is really necessary for the performance of the contract or if it is really necessary to take steps at the request of the data subject prior to entering into a contract. Both consumer law and data protection law refer to consent and contracts in all facets of user interaction within the data economy. Accordingly, this project aims to create a theoretical framework that would allow for a common understanding of the concept of consent in data protection and in contract law that would enable the legitimate processing of personal data and the ability to enter into a contract more consciously.

## Project 7 – The Value of Human Autonomy in Legal AI

This project looks at the changing role and position of human beings in relation to AI technologies. In particular, it focusses on the legal and ethical meaning of human autonomy in judicial systems pervaded by AI technologies. The starting point is that AI applications are not neutral instruments but have a fundamental impact on what it means to be human. As data-driven decision-making tools are increasingly designed to act and even proactively intervene in seemingly autonomous ways – without any human beings in the loop – this immediately evokes the fear that these technologies will 'take over' and diminish human autonomy and discretion. Accordingly, this project examines and re-evaluates human autonomy in the data-driven regulatory paradigm.

Rather than framing AI as an unstoppable force that will dominate and overrule human autonomy, this research assesses the possibility of understanding human autonomy as inherently relational and connected to AI.

This project focuses on how the use of AI in the judicial system affects and interacts with the autonomy of both citizens and public authorities, such as judges and prosecutors. How does AI shape human discretion in the judicial system? How is responsibility allocated in relation to systems that may be opaque to those who work with them and are usually constructed by private-sector actors? How does the shift to a data-driven paradigm shape the positioning of human autonomy in conceptions and evaluations of the rule of law and associated problems, such as due process, predictability and fairness? Consequently, the objective of this project is to examine and re-evaluate the value of human autonomy in the shift towards a data-driven regulatory regime in order to identify possible ways of safeguarding this key value in legal systems.

### Project 8 – Protecting Autonomy in a Data-Driven Society

The freedom of individuals to make informed and uncoerced choices is at the heart of modern democracies. However, digital giants are gaining an increasing level of control over individuals. By tracking our behaviour and using data analytics tools, these companies know more about our interests than we do ourselves, thereby enabling them to manipulate our preferences. These concerns no longer only relate to the simple consumption of goods and services; they also affect our autonomy in the consumption of news and the beliefs that ground political voting behaviour. Digital giants are furthermore turning into gatekeepers that other businesses depend on for access to markets and to reach consumers. Through the design of their platforms, digital giants also determine how individual and business users can express themselves and innovate on the basis of the tools provided. Digital giants are thereby able to arbitrarily impose their own rules on the competitive process and the autonomy of businesses and individuals, as well as on the creativity and freedom of expression of others.

Competition, consumer and intellectual property law are becoming increasingly relevant as regimes that can provide additional mechanisms to protect autonomy in a data-driven society. On the one hand, the way in which behavioural targeting and personalisation restrict the freedom of choice of individuals also affects the nature of competition and thereby triggers competition law issues. This changes the relationship between competition and consumer law. While competition law is traditionally concerned with protecting the availability of a range of consumer options through competition, consumer law is primarily responsible for protecting consumers' ability to choose effectively among these options. However, if a consumer is nudged towards one option by having their ability to choose freely restricted, the effect may be the same as only

having one option, so that competition issues occur. On the other hand, the 'gatekeeping' nature of platforms and the resulting dependence of other businesses on them raises questions about the role of level playing field discussions in competition and internal market law more broadly. There is therefore a need to take a more holistic view of the different regimes that are at stake in the protection of the autonomy of individuals and businesses in a data-driven society, which this project lays the groundwork for.

## Project 9 – Evaluating Code as a Mechanism for Regulating AI in Healthcare

The final project starts from the premise that the range of emerging and potential applications of AI in health and medicine is substantial, offering promise for more efficient and effective healthcare, as well as undergirding the shift to personalised medicine in its preventive, therapeutic and care dimensions. Already, various forms of machine learning are improving on or equalling physicians' performance in diagnosis (e.g. diabetic retinopathy and skin lesions), and additional complex clinical processes are currently under investigation (e.g. prognosis, prevention and treatment decisions). However, the healthcare domain operates within the context of strictly observed normative frameworks that range from fundamental rights to ethical and professional norms that are deeply entrenched. Some of these algorithmic applications run up against legal, ethical, policy and social aspects of these normative frameworks. For example, data protection law specifically deals with issues of privacy and protection of personal data but also with issues such as transparency and non-discrimination (e.g. automated decision-making). These in turn may invoke other legal, ethical and social considerations through the impact that the use of AI systems may have on clinical and care practices.

AI offers opportunities in all phases of the research-clinical care spectrum (diagnosis, prognosis and treatment), as well as in the context of social work and prevention. However, the existing regulatory schemes that target human behaviours demonstrate gaps in the ability to reach aspects of the use of AI in medicine and health that could violate existing norms, and the values and interests that these norms are designed to promote and protect. For example, prognoses based on algorithms have the potential to trigger several concerns. Are there ways to facilitate 'non-discrimination or equal access by design' or to promote preservation of key aspects of the doctor-patient relationship or shared decision-making by the use of code? Additionally, data collection for research or clinical care may offer opportunities for code in the service of non-discrimination or data minimisation/protection. Thus, this project examines the interaction of healthcare personnel with AI systems in the healthcare and social work context. It also explores the opportunities and merits of the use of 'code' in AI as a way of translating key aspects of existing normative frameworks that govern the health domain. This project examines

where code may be appropriate based on legal, ethical and practical criteria, what options for code may look like, and the merits of using code versus alternative regulatory measures or modalities, including existing governance mechanisms.

## 5. This Book's Contents

This book contains the first output of the nine projects. As such, it contains a wide range of topics, ideas, questions and dilemmas. Part II contains five shorter chapters that show how classic divides that underlie many of the contemporary legal frameworks, in particular the divides between the private and the public domain, between the private and the public sector, and between private and public law, are challenged in the data-driven environment. Given these complexities, new legal frameworks have been proposed and adopted by the European Union, the Council of Europe and other organisations with regulatory powers. Notable regulation includes the GDPR, the Digital Services Act, the Digital Markets Act, the AI Act and the AI Liability Directive. There is discussion, however, regarding to what extent these laws are capable of adequately addressing the many challenges that arise in the data-driven environment. Part III discusses the revision of doctrines that are typically associated with private sector bodies, such as consent, consumer law and competition law, while Part IV assesses doctrines that find their origin in public law, such as fairness, non-discrimination and justice. Part V concludes.

### Part II – Challenges to the Private-Public Divide

In Chapter 2, Shweta Degalahal shows that the groundwork for most theories on democracy, the rule of law and public participation depends on the conceptualisation of a public domain. In the digital sphere, however, a thorough conceptualisation is lacking. This chapter shows that there are many trends that undercut the traditional idea of the public domain in the digital realm. In particular, it discusses how the public space can be reconceptualised in light of a right to informational privacy.

In Chapter 3, Manos Roussos builds on a similar idea, namely the classic divide between the public and the private sector, where public sector organisations execute public sector tasks and private sector organisations pursue private interests. This model, however, no longer works in the digital era. Banks, for example, are required by law to help with tracing terrorist financial transaction and to identify signs of money laundering in order to aid law enforcement authorities in their tasks. To complicate matters further, combating money laundering and terrorist financial transactions requires data sharing between parties in different jurisdictions. This means that often, four or more legal instruments apply (e.g. the EU framework for data processing by private parties as laid out in the GDPR, the EU framework for data processing by law enforcement authorities as laid out in the Law Enforcement

Directive, and the US frameworks for data processing by banks and by law enforcement authorities). Ensuring these are applied effectively and legitimately presents a challenge.

Chapter 4, written by Taner Kuru, explores the erosion of the public-private divide by homing in on the use by law enforcement authorities of DNA databanks held by private sector organisations. Although this has led to the resolution of several cold cases, there are obvious concerns of privacy. The legal regime is inadequate in providing proper safeguards and questions arise as to the legitimacy of the actions by the police as European courts have adopted several judgments that do not yield clear and detailed guidelines.

Both Keyomars Khaleghi, in Chapter 5, and Aimen Taimur, in Chapter 6, discuss the importance of microtargeting, profiling and manipulation in the area of elections. Through the use of these tactics, as well as by using fake news, both private parties and foreign agents can affect the outcome of political processes, elections and decision-making. Khaleghi discusses these developments in light of the right to free elections and assesses to what extent the contemporary legal framework needs reconfiguration. Taimur, in turn, asks how we can protect our mental integrity and *forum internum* through a rights-based approach.

## Part III – Reconfiguring the Legal Paradigm for the Private Sector

Ana-Maria Hriscu and Eleni Kosta present their thinking on the role of consent in the data-driven era in Chapter 7. They show that informed consent, as a legal basis for processing personal data, has been framed in the EU data protection law framework as a tool to empower individuals. However, the analysis of processing activities of three very large online platforms, Google, Meta and Microsoft, reveals how they fall short of meeting some of the main transparency and consent requirements set in the law. Therefore, in the online context, consent can be said to have a disempowering effect rather than an empowering one. Hriscu and Kosta discuss possible solutions to this problem.

In Chapter 8, Thomas Tombal and Inge Graef analyse whether the Data Act can reach its objective of stimulating the European data economy. They show that the Data Act is a welcome but complicated piece of legislation that brings together various interests, in particular the interest of the data holder in protecting its investments, the interest of third parties in accessing data to develop own products and services, and the interest of individuals in protecting and controlling the use of their personal data. This chapter provides a detailed discussion of the Data Act, including critical reflections on how the Act is framed.

Competition law and merger regulation, in particular in the healthcare sector, is the focus of Chapter 9. In this contribution, Tjaša Petročnik and Inge Graef provide a review of data-driven mergers in the healthcare sector that illustrates how economic concerns relating to competition become increasingly intertwined with non-economic considerations regarding privacy and health protection, among other aspects. The chapter argues that this development requires a more proactive approach by the European Commission

and national authorities, and it makes suggestions for achieving better coordination and cooperation within the existing regulatory framework.

### Part IV – Reconfiguring the Legal Paradigm for the Public Sector

In Chapter 10, Aviva de Groot and Siddharth Peter de Souza argue that the regulation of AI by the EU lawmaker provides a particular momentum for critical engagement. There is broadly voiced urgency to legally protect against AI-fuelled harms but, they argue, the digital rights field needs to engage with those harms from a less privileged standpoint, in order to ensure that society is just and fair. They map justice-related aspects of automated decision-making and argue that these should be considered when updating and revising the AI Act.

Bart van der Sloot, Merel Noorman and Linnet Taylor discuss the theoretical framework that guides anti-discrimination law. They argue that on many accounts, this framework needs a reconceptualization in light of the rise of AI. The current paradigm is ill-applicable to the potential arbitrary ways in which AI can make decisions as well as reproduce and reinforce data biases. This is why they propose understanding discrimination law's main rationale as requiring adequate and sufficient reasons for making decisions, rather than preventing decisions made on the basis of racial, religious, sexual or other distinctions. In Chapter 11, they explore the possibility of drawing inspiration from the republican idea of freedom, which is not based on being free from interference, as is the liberal conception of freedom, but on non-domination.

In Chapter 12, Bart van der Sloot discusses legal AI, or the ideal of automating law-making and in particular, assisting or replacing judicial decisions in which laws are applied to specific cases. The chapter shows that there are three prominent philosophies of law: the legal positivist view of the legal order, the natural law theory and legal pragmatism. The ideal of legal automation aligns with the legal positivist view while, if law is understood through the lens of natural law theory, it may be complicated to capture the essence of the legal order in code. For legal pragmatists, however, the very idea of legal automation goes against what it means to have a legal order. Van der Sloot explains this point by discussing the philosophical framework of Lon Fuller in detail.

### Part V – Conclusion

Finally, Chapter 13 contains the concluding chapter and provides an overview of the most important lessons drawn from this book as well as an outlook to future research necessary to answer the many questions, dilemmas and intricacies raised by the authors of the various chapters.

PART

# II

# CHALLENGES TO THE PRIVATE-PUBLIC DIVIDE

CHAPTER

# II

# Informational Privacy Rights in the Digital Public Space

**Shweta Reddy Degalahal**[1]

https://doi.org/10.26116/azyc-zk26

1 PhD Researcher, Tilburg Institute of Law, Technology, and the Society (TILT), Tilburg Law School (TLS), Tilburg University

## 1. Background

In 2020, reports of Clearview AI designing and deploying a facial recognition tool that was trained using publicly available images from social networking sites started to surface. Clearview AI used a data collecting technique called 'web scraping'. This is done using automated software to extract large amounts of information from the internet or a specific digital platform.[2] According to their privacy policy, additional information – such as geolocation metadata and other information that can be derived from facial appearances – from publicly available images was also being collected.[3] This data collection was not done using a targeted approach. Instead, it was done by collecting and processing publicly available images regardless of users' jurisdiction.

The EU's General Data Protection Regulation (hereinafter 'GDPR' or 'Regulation') does not provide for exemptions from its application if personal data are publicly available. On the contrary, 'publicly available' is not defined in the Regulation nor is there a general authorisation for the reuse of such personal data. This implies that the scope of obligations prescribed under the Regulation are not dependent on the accessibility or availability of personal data but extend to any and all personal data processing operations unless specifically exempted. Hence, it was not a surprise when multiple data protection authorities held that these data processing operations were in violation of the GDPR; specifically the provisions related to lawful ground of processing personal data, data storage limitation and the information obligations.[4] Considering that the scope of the obligations extends to publicly available personal data, the next step is to examine if these obligations are sufficiently adequate to safeguard personal data in the digital public space. This chapter briefly examines areas of further research, by conceptualising the notion of 'public' for the digital space and evidence of privacy harms that arise due

2 Zamora, A. (2019). Making room for big data: web scraping and an affirmative right to access publicly available information online. *J. Bus. Entrepreneurship & L.*, *12*, 203.

3 *Restricted Committee Deliberation No° SAN-2022-019 of 17 October 2022 concerning CLEARVIEW AI,* Commission Nationale Informatique & Libertes (17 October 2022), <https://www.cnil.fr/sites/default/files/atoms/files/deliberation_of_the_restricted_committee_no_san-2022-019_of_17_october_2022_concerning_clearview_ai.pdf>.

4 Information Commissioner's Office, *ICO issues provision view to fine Clearview AI*, Information Commissioner's Office (29 November 2021) <https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2021/11/ico-issues-provisional-view-to-fine-clearview-ai-inc-over-17-million/>. Hamburg Commissioner for data protection and freedom of information, *Consultation prior to an order pursuant to article 58(2)(g) GDPR*, NOYB (27 January 2021) <https://noyb.eu/sites/default/files/202101/545_2020_Anhörung_CVAI_ENG_Redacted.PDF>. Garante Per La Protezione Dei Dati Personali, *Injunction order against Clearview AI*, (10 February 2022) <https://www.gpdp.it:443/web/guest/home/docweb/-/docweb-display/docweb/9751362>. Hellenic Data Protection Authority, *Επιβολή προστίμου στην εταιρεία Clearview AI, Inc* (13 July 2022) <http://www.dpa.gr/el/enimerwtiko/prakseisArxis/epiboli-prostimoy-stin-etaireia-clearview-ai-inc>.

to the unauthorised use of publicly available personal data. This is done to determine whether the General Data Protection Regulation is effective in safeguarding personal data that is part of the digital public space.

One of the strongest obligations under the GDPR in favour of the data subjects is the requirement to provide a privacy notice (i.e. the information obligations). This ensures that the data subject is aware of the processing operations and their rights associated with such processing. For collecting publicly available personal data, entities have to rely on the information obligations set forth in Article 14 of the GDPR. This Article provides details on the information to be provided to the data subject where personal data has not been directly obtained from them. This information includes the categories of collected personal data, information regarding their data subject rights, the source of personal data and so forth.[5] However, implementation poses a considerable challenge.

This challenge can be examined through Clearview AI's scraping activities on Facebook. Users share their data with Facebook, meaning that for the purpose of the transaction, Facebook is the data controller, and the user can be expected to read Facebook's privacy policies to understand the details of how their information is processed. Clearview AI is not a data processor of Facebook; Facebook has not outsourced any data processing operations to Clearview AI. Since it's not a data processor, Facebook is not obligated to include Clearview AI's web scraping activities in their privacy policy. Yet due to the indiscriminate data collection process of web scraping, Clearview AI does not have a direct relationship with users.[6] There does not seem to be an efficient method to isolate users whose data has been extracted and to contact them and provide relevant information under the GDPR. It is unclear how Clearview AI is expected to initiate processing operations in this case because there is no reasonable method for providing Facebook users a privacy notice prior to the extraction of the data. Hence, even though the safeguards prescribed in GDPR extend to personal data that can be considered publicly available, the practical considerations of implementing these obligations need to be revisited and the extent of the safeguards provided needs to be examined. An argument can be made here suggesting that Clearview AI's inability to legally process personal data through scraping by providing an adequate privacy notice is evidence of the legal system protecting publicly available personal data. However, it took journalistic efforts to trace the workings of Clearview AI, and this then led to data protection authorities initiating an investigation into their operations to indicate an enforcement issue. Still, the larger concern remains: due to the nature of publicly available personal

5 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, Article 14.

6 Zamora, A. (2019). Making room for big data: web scraping and an affirmative right to access publicly available information online. *J. Bus. Entrepreneurship & L.*, *12*, 203, p.10

data, there is no regulatory friction preventing organisations like Clearview AI from collecting such data at a large scale. This leads to a potential loss of informational privacy due to lack of control over the further usage of personal data. This in turn warrants conversations about ex-ante protections for publicly available personal data at the platform level that does not rely solely on journalists' ability to unearth wrongdoing.

Clearview AI is not the only incident of publicly available personal data being reused through web scraping. ChatGPT is a large language model[7] that was trained using text-based databases scraped from the internet. This model was trained to translate, predict and generate text in response to user queries.[8] Details about the datasets used to train the model have been classified as proprietary. However, the Italian data protection authority issued an emergency decision,[9] ordering Open AI to stop using personal information in the training data. Previous iterations were criticised as posing a privacy risk because, when queried repeatedly, they could respond with personal information about individuals who were not public personalities.[10] Answers to questions like what threshold was used to determine who amounts to a public personality are not clear due to the proprietary nature of the training datasets and other processing information. Law enforcement agencies have also relied on publicly available social media communications for investigating, prosecuting and preventing criminal activities.[11] Additionally, web scraping is used by researchers as a technique to collect and analyse large scale datasets from social media for research purposes.[12]

The disclosure of personal data to the general public may also be mandated by statutory law. For example, to prevent conflicts of interest and corruption in the public sector, public servants and their spouses or partners are required to disclosure their personal data under Lithuania's law on the reconciliation of public and private interests in public service.[13] The

7 Lee, A. (2023) *What Are Large Language Models Used For and Why Are They Important?* NVIDIA Blog (26 January 2023), <https://blogs.nvidia.com/blog/2023/01/26/what-are-large-language-models-used-for/>.

8 Ruby, M. (2023) *How ChatGPT Works: The Models Behind The Bot*, Medium (31 January 2023) <https://towardsdatascience.com/how-chatgpt-works-the-models-behind-the-bot-1ce5fca96286>.

9 Garante Per La Protezione Dei Dati Personali, *Provision of March 30 2023 doc. web no. 9870832,* (30 March 2023) <https://www.gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9870832>.

10 Melissa Heikkilä, *What does GPT 3 "know" about me?,* MIT Technology Review (31 August 2022) <https://www.technologyreview.com/2022/08/31/1058800/what-does-gpt-3-know-about-me/>.

11 Lilian Edwards and Lachlan Urquhart, *Privacy in Public Spaces: What Expectations of Privacy Do We Have in Social Media Intelligence?* International Journal of Law and Information Technology 24(3). 279,310 (2016).

12 Kusumasari, B., & Prabowo, N. P. A. (2020). Scraping social media data for disaster communication: how the pattern of Twitter users affects disasters in Asia and the Pacific. *Natural Hazards*, *103* (3), 3415-3435.

13 Lietuvos Respublikos viešųjų ir privačių interesų derinimo valstybinėje tarnyboje įstatymas Nr. VIII-371 (Law No VIII-371 of the Republic of Lithuania on the reconciliation of public and private interests in the public service) of 2 July 1997 (Žin., 1997, No 67-1659).

requirement has been challenged on account of its interference with the privacy of public servants and their family members. However, the court reasoned[14] that the objective of ensuring transparency in public service was proportionate to warrant the interference of privacy despite the argument that the cumulative effect of personal data attributes being disclosed could reveal sensitive details of the private lives of the public servants.

The unauthorised use of publicly available personal data is not limited to large private corporations. Users of dating applications have taken screenshots of other users on the platform that they then disclosed to a larger audience.[15] A journalist posing as gay man on Grindr flew to the Olympic village in 2016 and published identifying information of closeted Olympians using Grindr on their news website; this was subsequently taken down after collective outrage.[16] This example can be distinguished from the more recent fines against Grindr by the Norwegian Data Protection authority.[17] In the latter, Grindr violated data protection legislation by disclosing user data to a third party providing advertising services. In the former, it was a user of the platform that violated the privacy of the other users under the pretext of it being a public interest story.

Based on these limited examples, it can be inferred that any legal safeguard for protecting publicly available personal data has to be effective regardless of who the alleged infringer is. In other words, infringements can occur in private entity to user, user to user and state to user relationships. However, prior to examining the efficacy of safeguards for publicly available personal data, a conceptual approach towards understanding the interaction of the broader right to informational privacy in the digital public space is necessary.

## 2. Need for Defining a Digital Public Space for the Purposes of Informational Privacy

Public spaces are viewed as open, uncontrolled spaces where the ability of an individual to exert control over disclosure of private information is minimal. Private spaces, on the other hand, are secluded, controlled spaces where individuals are able to exert control over the disclosure of their private information.[18] This division between spaces also leads

14 C – 184/20, *OT v Vyriausioji tarnybinės etikos komisija* ECLI:EU:C:2022:601.

15 Cobb, C., & Kohno, T. (2017, April). How public is my private life? Privacy in online dating. In *Proceedings of the 26th International Conference on World Wide Web* (pp. 1231-1240).

16 Mcnamara, B. (2016). *Olympics 2016: Closeted Gay Athletes Outed by Daily Beast Grindr Article*, Teen Vogue (12 August 2016) <https://www.teenvogue.com/story/straight-journalist-outed-closeted-gay-olympic-athletes-grindr>.

17 European Data Protection Board, *Norwegian DPA Imposes Fine against Grindr LLC* (21 December 2021) <https://edpb.europa.eu/news/national-news/2021/norwegian-dpa-imposes-fine-against-grindr-llc_en>.

18 De Souza e Silva, A., & Frith, J. (2010). Locational privacy in public spaces: Media discourses on

to a spatial conceptualisation of informational privacy as pertaining to personal facts that belong to the private space of the individual as opposed to the public space, which contains public facts.[19] There has been a rise in surveillance of physical public places. This has occurred due to the deployment of cameras capable of facial recognition by the state, the introduction of lateral surveillance mechanisms by neighbourhood associations, products like Google Glass, drones and the push towards smart cities. Consequently, there is a growing consensus that the impact of such measures and products on informational privacy in public needs to be examined.[20] However, the examples of web scraping and the disclosure of other personal data also indicates the rise of surveillance on digital platforms that provide widely accessible personal data.

The current literature does not point towards a consistent definition of a digital public space. Understandings of what amounts to a digital public space include spaces[21] wherein personal information is hypothetically or freely accessible,[22] a space to engage in political participation,[23] a space that promotes freedom of speech and expression and much more. With specific reference to data protection, Article 9 GDPR prohibits the processing of special categories of personal data unless specific exceptions apply; one of these is personal data that has been 'manifestly made public by the data subject'. The regulation does not provide a definition for this phrase. However, in its guidance[24] on targeting social media users' data, the European Data Protection Board (EDPB) states that such a determination should include an analysis of the default private settings of the data subject, the nature of the social media platform, and information provided on the data subject regarding the public nature of the disclosed information.[25] It is essential to note that this determination is only for the special categories of personal data under GDPR and for personal data revealing such special categories as defined in the Regulation. It does not extend to personal data as in general. This determination is also specif-

location-aware mobile technologies. *Communication, Culture & Critique*, *3* (4), 503-525.

19 Brincker, M. (2017). Privacy in public and the contextual conditions of agency. In *Privacy in public space* (pp. 64-90). Edward Elgar Publishing.

20 Madiega, T. & Mildebrath , H. *Regulating facial recognition in EU,* European Parliamentary Research Service (16 September 2021) https://epthinktank.eu/2021/09/16/regulating-facial-recognition-in-the-eu/ 'An Act To Increase Privacy and Security by Regulating the Use of Facial Surveillance Systems by Departments, Public Employees and Public Officials', State of Maine.

21 This chapter will use space and space interchangeably and does not delve into the differences between the both of them as laid out in social sciences theory.

22 Hartzog, W. (2019). The Public Information Fallacy. *BUL Rev.*, *99*, 459.

23 Koops, B. J., & Galič, M. (2017). Conceptualizing space and place: Lessons from geography for the debate on privacy in public. *Privacy in public space*, 19-46.

24 European Data Protection Board, *Guidelines 08/2020 on the targeting of social media users*, European Data Protection Board (13 April 2021) <https://edpb.europa.eu/system/files/2021-04/edpb_guidelines_082020_on_the_targeting_of_social_media_users_en.pdf>.

25 Ibid.

ically limited to social media platforms that provide individuals with the option to control their privacy settings and on which individuals intentionally disclose personal data. It does not account for unintentional disclosure of personal data, illegal disclosure of data (data breaches), coerced disclosure due to lack of choice and so forth. Considering that EDPB guidance is limited to special categories of personal data and does not extensively analyse what public means for the digital space, further research may be necessary. This research might focus on what accounts for public in the digital space by taking into consideration the different conceptualisations of privacy, as well as the privacy harms that can arise in such a space.

## 2.1 Existing Conceptualisations of Privacy in Relation to the Public Space

It has been argued that defining privacy by locating its importance in achieving other freedoms and rights of the individual can fail to include the privacy expectations that an individual might have. This is because privacy is dependent on the dynamic relationship between the individual and the social and cultural contexts they interact with.[26] Privacy can be conceptualised in terms of access, namely a zone of privacy around the individual that is inaccessible to society at large or that permits restricted accessibility. However, this has spatial undertones as there is a presumption of a boundary that prevents access. It has been argued that these boundaries are managed by individuals in accordance with the types of relationships people maintain, and they may fluctuate over time.[27] If the expectation of informational privacy is dependent on whether someone has access to information about the individual, the very fact that disclosing information in a specific context provides access to multiple individuals can mean an unavoidable loss of privacy.[28] However, the conceptualisation of privacy in terms of control (i.e. of the individual being in control over who has access to their information) aids in enabling privacy in the public space.

It has been argued that, instead of defining privacy in opposition to the public space, it is useful to define privacy in functional terms (i.e. what privacy is expected to achieve).[29] It is here that Westin's definitions of four basic states of privacy[30] in an increasing level of the individual's involvement in the public space can be useful. These

26 Cohen, J. E. (2012). What privacy is for. *Harvard Law Review*, *126*, 1904.
27 Koops, B. J. (2018). Privacy Spaces. *West Virginia Law Review*, *121*, 611.
28 Gavison, R. (1980). Privacy and the Limits of Law. *The Yale law journal*, *89* (3), 421-471.
29 Suresh, H., & Guttag, J. (2021, October). A framework for understanding sources of harm throughout the machine learning life cycle. In *Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization* (pp. 1-9).
30 Westin, A. F. (Ed.). (1971). *Information technology in a democracy*. Harvard University Press.

four states are solitude, intimacy, anonymity and reserve.[31] In the context of understanding privacy in the digital public space, the states of anonymity and reserve that he defines are specifically relevant. He defines anonymity as a state where the individual in a public place still seeks to find freedom from identification and surveillance; the individual desires 'public privacy'. By 'public privacy', he means that the individual does not expect to be personally identifiable and held to the full rules of expected social behaviour by the observers. According to Westin,[32] the state of reserve involves the creation of a psychological barrier against unwanted intrusions; it is the individual's choice whether or not to disclose information to the people surrounding them. He argues that the ability to preserve an individual's choice of disclosure refers to the dynamic aspect of privacy in daily interpersonal relations.

This dynamic aspect of privacy is captured by the contextual integrity (CI) framework proposed by Helen Nissenbaum. This framework rejects the traditional public/private dichotomy. Instead, it proposes an alternate concept of privacy based on the information norms pertaining to each context of disclosure.[33] She argues that there are multiple contexts in which individuals disclose information, and these contexts are governed in accordance with distinctive norms pertaining to each context.[34] The expectations of privacy that individuals have are in accordance with the norms that govern the contexts in which information is disclosed. The framework provides for a decision heuristic[35] that considers the information flows, prevailing context, transmission principles, entrenched information norms and other factors to determine whether a new practice or new processing operation is in violation of the principles of the framework. For example, an application that collects multiple data types can have distinct context-specific norms because the privacy expectations that people hold over different information also differs. This points to the fluid dimensions of privacy, which can extend to personal information that is considered publicly available because of the mere accessibility of that information.

The contextual integrity framework has been criticised for its reliance on entrenched informational norms derived from the societal norms prevalent at the time of analysis.[36] The framework does not consider the possibility that the entrenched informational norms against which legitimacy of new practices are evaluated can undermine the conditions of privacy and democracy by their very nature. For example, in the Clearview AI case, it can be argued that, due to the ever increasing cases of ambient surveillance,

31 Ibid.
32 Ibid.
33 Nissenbaum, H. (2009). Privacy in Context: Technology, Policy, and the Integrity of Social Life.
34 Ibid, p. 300.
35 Ibid, p. 379.
36 Reidenberg, J. R. (2014). Privacy in public. *U. Miami L. Rev.*, *69*, 141.

an individual should expect the possibility of a web scraper collecting and processing their personal data. The analysis of whether the ambient surveillance, which can be considered as a norm, is a violation of the privacy rights of the individual is not necessarily considered by the CI framework. Joel Reidenberg provides for an alternative framework that moves away from the reasonable expectation of privacy of the individual.[37] He suggests distinguishing the true public space of governance from private actions, and he only seeks protection for the latter. He argues that governance related acts include activities that have an inherent public significance based on the social norms around what constitutes public significance and public interest.[38] By default, acts that are not governance-related are private acts in the public space and need to be protected. Yet despite his criticisms of the contextual integrity framework's reliance on social norms to determine validity of a practice, he relies on social norms to determine what amounts to acts of public significance, hence giving rise to the same challenges.

While the EU data protection framework is this chapter's primary focus, the conceptualisations of privacy mentioned above are predominantly from American scholars. These were chosen to be part of this brief analysis as they were the most prominent theories that specifically discussed privacy implications of actions in the public space. These different conceptualisations highlight the necessity of preserving the dynamic contexts of specific privacy expectations in the digital public space in order to effectively safeguard associated rights and freedoms in a democratic society. In doing so, it should be noted that overreliance on existing informational norms or widely practiced business models (e.g. of web scrapers and associated data markets) without analysis of their legitimacy will not actually contribute to protecting individuals' right to privacy in the digital public space.

## 2.2 Typology of Privacy Harms in the Digital Public Space

In addition to an analysis of privacy in the digital public space, attention has to also be diverted to the privacy harms that can occur in such a space. Existing literature on privacy harms broadly categorises them into physical, economic, reputational, psychological, autonomy-based, discrimination and relationship harms.[39] Based on the definition of the digital public space that will be identified, it is possible that the potential privacy harms discovered will fit into this existing typology. However, in the absence of such a definition at this stage, potential privacy harms can be deliberated on. These are based on the Clearview AI case in which the scraped personal data was used to train facial

37 Ibid.
38 Ibid.
39 Keats Citron, D., & Solove, D. J. (2022). Privacy Harms, 102 Boston Univ. *L. Rev*, *793*.

recognition algorithms, which were then sold to law enforcement officials.[40] Privacy harms can occur at two stages: at the stage of collection and at the stage of further use when the facial recognition algorithm is deployed.

At the collection stage, individuals who disclosed their personal data were unaware of the activities of the web scraper that could give rise to autonomy-based harms. As such, the individual's ability to assess the risk of disclosing personal data on a platform where it can be collected using a web scraper was affected. It is acknowledged that the harmful and possibly far-reaching effects of web scraping by Clearview AI may not be immediately apparent. However, it is essential that any further assessment of privacy harms in the digital public space considers the risk of future harms that may arise due to such an incident.

At the stage of further use, when facial recognition algorithms are deployed by law enforcement authorities, potential discrimination[41] propagated by such algorithms could be considered a privacy harm. However, the same issue with causality arises here as well. There is a need for extensive empirical research on the privacy expectations of individuals in the digital public space to supplement the literature on privacy harms in general. There is also a need to determine the factors that determine the causality between the privacy harms identified and the digital public space from where the publicly available personal data was initially collected.

## 3. Conclusion and Further Questions

The literature on privacy in physical public spaces was helpful in advocating for privacy and data protection rights of the individuals when digital technologies started encroaching the physical public spaces. However, there is neither a definite theory of informational privacy that extends to the digital public space nor a consensus on the contours of a digital public space. Private entities, law enforcement agencies and research organisations have increasingly begun to rely on personal data disclosed on digital platforms for their own purposes. This may be a strong indication of the need to revisit the notion of data protection safeguards based on the public and private nature of data disclosure by individuals. Existing safeguards under the GDPR for personal data that could be deemed public may be ineffective either due to implemen-

40 Hill, K. (2020). *The Secretive Company that might end privacy as we know it,* The New York Times (18 January 2020) <https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html>.

41 Crockford, K. (2020). *How Is Face Recognition Surveillance Technology Racist?*, American Civil Liberties Union (16 June 2020) <https://www.aclu.org/news/privacy-technology/how-is-face-recognition-surveillance-technology-racist>

tation challenges or due to protection being limited to special categories of personal data.

The GDPR's degree of ineffectiveness in protecting publicly available personal data is unclear at the moment. There is a need for empirical evidence of individuals' ability to safeguard their publicly-available personal data that takes privacy expectations into consideration. The lack of any ex-ante data protection authority investigations and any explicit mention of publicly available personal data does provide for a prima facie reason to conduct such empirical research. Further research into the phrases 'public' and 'private' etymologically and from social science theory and political theory perspectives, and from ethnographic accounts in urban governance literature could provide insight into the importance and characteristics of public spaces – the physical ones, at least. The difference between physical public spaces and digital public spaces, owing to an individual's inability to participate in the latter without leaving digital trails and a lack of conscience anonymity and obscurity, suggests that current research on protecting privacy and data protection in physical public spaces cannot be directly translated to digital spaces. The insights derived from the research should thus be adapted to the digital environment first. Based on the conceptualisation of a digital public space and the empirical evidence on individuals' expectations of privacy in the digital space, the current data protection framework in the EU needs to be evaluated to examine the need for any potential additional safeguards.

CHAPTER

# III

# Transatlantic Personal Data Transfers Between the EU and the USA for Anti-Money Laundering and Countering the Financing of Terrorism Purposes

**Manos Roussos**[1]

https://doi.org/10.26116/32zt-ca46

1 PhD Researcher, Tilburg Institute of Law, Technology, and the Society (TILT), Tilburg Law School (TLS), Tilburg University.

## 1. Introduction

In today's society and economy, data is increasingly considered to be a new form of commodity[2] commonly referred to as 'the new oil'.[3] This change has taken place as the data economy initially developed as a result of the growth of data technology and big data being in a constant state of change.[4] This has led to continuously emerging challenges to the contemporary legal system owing to economic and technological advancements. Industries and working environments that are driven and characterised by data processing and international data transfers of data on a massive scale[5] include healthcare,[6] transportation[7] and banking.[8] Data flows between related parties are massive in size and quite frequent,[9] while large volumes of data are processed, stored and analysed in real-time or near real-time. The environments in which data are processed in such a manner are often referred to as 'data-intensive environments'.[10] In these environments, a large amount of the processed data constitutes personal data, namely information that relates to an identified or identifiable living individual.[11] This results in companies and organisations that conduct business in these industries creating large databases containing personal data.

2 De Franceschi, A., & Lehmann, M. (2015). Data as tradeable commodity and new measures for their protection. *Italian LJ*, *1*, 51. Zelianin, A. (2022). Personal Data as a Market Commodity in the GDPR Era: A Systematic Review of Social and Economic Aspects. *Acta Informatica Pragensia*, *11* (1), 123-140.

3 The Economist, *The world's most valuable resource is no longer oil, but data* (6 May 2017) , <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>.

4 Zelianin, A. (2022). Personal Data as a Market Commodity in the GDPR Era: A Systematic Review of Social and Economic Aspects. *Acta Informatica Pragensia*, *11* (1), 123-140.

5 OECD, *Data-driven Innovation for Growth and Well-being*, 2014. <https://www.oecd.org/sti/data-driven-innovation-9789264229358-en.htm>.

6 Muller, S. H., Kalkman, S., van Thiel, G. J., Mostert, M., & van Delden, J. J. (2021). The social licence for data-intensive health research: towards co-creation, public value and trust. *BMC Medical Ethics*, *22* (1), 110.

7 Ngo, L. B. (2017). Data Science Tools and Techniques to Support Data Analytics in Transportation Applications. In Chwodhury, M., Apon, A. & Dey, K. (eds.) *Data Analytics for Intelligent Transportation Systems*, Elsevier.

8 Aitken, M., Ng, M., Horsfall, D., Coopamootoo, K. P., van Moorsel, A., & Elliott, K. (2021). In pursuit of socially-minded data-intensive innovation in banking: A focus group study of public expectations of digital innovation in banking. *Technology in Society*, *66*, 101666.

9 Aımeur, E., Brassard, G., & Guo, M. (2022). How data brokers endanger privacy. Transactions on Data Privacy, 15(2), 41-85.

10 Middleton, A. M. (2010). Data-intensive technologies for cloud computing. *Handbook of cloud computing*, 83-136.

11 Article 4(1) 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), OJ L 119, 4.5.2016, p. 1–88.

The richness of information contained in such data pools is becoming more and more useful for law enforcement authorities (LEAs) as a supportive mechanism to combat criminal activities.[12] Data processing plays a crucial role in that regard by providing them with the ability to analyse and understand criminal activity more effectively.[13] By collecting and processing data from various sources, LEAs are able to identify patterns and behaviours that can help them solve crimes, prevent future criminal activity and apprehend perpetrators. A prominent field in which such data pooled by private companies is quintessential for purposes and goals of LEAs is the combatting of financial crime,[14] in particular combatting money laundering and the financing of terrorism, two of the most widespread financial crimes globally. Money laundering is the procedure through which criminals 'clean' the benefits of their activities to hide their illegal origins. It is usually associated with the types of organised crime that generate huge profits in cash.[15] Following the definition given by Cox, it is the process of disguising the proceeds of illegal activities such as drug trafficking, embezzlement, fraud and other criminal activities as legitimate funds, for the purpose of concealing the origin, ownership or destination of the illegally obtained money or assets in order to avoid detection and prosecution by law enforcement agencies.[16] Countering the financing of terrorism is a core element in the fight against terrorism.

The financial sector is a prominent example of a data-intensive environment, where data are shared in a rapid, frequent and intense manner, supported by massive databases.[17] Especially in the financial sector, which is supported by modern technological advancements, financial institutions can analyse enormous amounts of data more quickly and effectively, and they may consequently comprehend and evaluate data by pooling them and using collaborative analytics.[18] This enables a more dynamic and

12 Europol, *Exploring Tomorrow's Organised Crime* (2015) <https://www.europol.europa.eu/sites/default/files/documents/Europol_OrgCrimeReport_web-final.pdf>.

13 Leese, M. (2023). Enacting criminal futures: data practices and crime prevention. *Policing and Society*, *33*(3), 333-347.

14 Deloitte, *The global framework for fighting financial crime | Enhancing effectiveness & improving outcomes,* The Institute of International Finance and Deloitte LLP White Paper (2019) <https://www2.deloitte.com/content/dam/Deloitte/tw/Documents/financial-services/tw-the-global-framework-for-fighting-financial-crime-en.pdf>.

15 Definition provided by the European Commission, Directorate-General for Migration and Home Affairs, <https://home-affairs.ec.europa.eu/policies/internal-security/organised-crime-and-human-trafficking/money-laundering_en>.

16 Cox, D. (2014). *Handbook of anti-money laundering*. John Wiley & Sons.

17 International Finance Corporation. (2019). Anti-Money-Laundering (AML) & Countering Financing of Terrorism (CFT) Risk Management in Emerging Market Banks.

18 FATF (2021), *Stocktake on Data Pooling, Collaborative Analytics and Data Protection*, <https://www.fatf-gafi.org/publications/digitaltransformation/documents/data-pooling-collaborativeanalytics protection.html>.

efficient identification of illegal activities, aiding the private sector and, consequently, the law enforcement sector to detect criminals and implement Anti-Money Laundering and Countering the Financing of Terrorism (AML/CFT) regulatory measures.[19] In this sense, banks and other financial institutions actively participate in managing money laundering and terrorism financing concerns.

In terms of territoriality, two of the main 'players' in the global landscape are the EU and the US. Personal data transfers between entities based in the territories of the two jurisdictions occur at a high frequency and volume for a range of purposes, including both commercial and law enforcement.[20] It is naturally expected that personal data are exchanged between entities in the financial sector too. The latter include transfers of personal data for AML/CFT purposes.

The aim of this chapter is to illustrate the unclarity of the current status quo regarding transfers of personal data between private and public/law enforcement entities and organisations established in the EU and the US for AML/CFT purposes. It seeks to do so by providing an overview of the existing legal regime and showcasing that there is an urgent need to clarify the landscape regarding personal data transfers between the EU and the US for AML/CFT purposes. This chapter is not meant to solve the problem. Rather, its main goal is to present the reader with an overview of the AML/CFT policy making landscape, to highlight the specific problem in relation to international data transfers between the EU and the US in the context of AML/CFT, and to argue for the urgent need for further research in this field.

## 2. AML/CFT Policymaking in a Nutshell

Although the practice of money laundering is not new,[21] AML/CFT has been a matter of high importance on a global scale throughout the past few decades. The main intergovernmental organisation to combat money laundering, terrorist financing and other related threats to the integrity of the international financial system is the Financial Action Task Force (FATF). FATF sets international standards and promotes the effective implementation of legal, regulatory and operational measures to combat money laundering. Its recommendations, commonly known as the 'FATF 40 recommendations', aim

19 International Finance Corporation. (2019). Anti-Money-Laundering (AML) & Countering Financing of Terrorism (CFT) Risk Management in Emerging Market Banks.

20 European Commission, Fact Sheet, *Digital Single Market – Communication on Exchanging and Protecting Personal Data in a Globalised World; Questions and Answers* (10 January 2017).

21 Some of the earliest known examples of money laundering date back to ancient China (as early as 2000 BC), when Chinese merchants would transfer their profits outside of the Chinese territory, in order to hide their wealth fearing that the local rulers would seize the merchants' profits. Morris-Cotterill, N. (2001). Money laundering. *Foreign Policy*, 16-22.

to provide a comprehensive framework for preventing, detecting and investigating money laundering and terrorist financing.[22] The USA and several EU Member states are FATF members. The global importance of AML/CFT can be noticed in the magnitude of the influence of global intergovernmental organisations like the FATF; it directly affects relevant policymaking across different jurisdictions. It can also be reflected in the timing of adoption of relevant European legislation. For instance, in 2003, the FATF revised its 40 recommendations, putting forward eight special recommendations on terrorist financing. Subsequently, in 2005, the EU legislator adopted an updated AML/CFT Directive (AMLD3),[23] that followed these updated recommendations. The current recommendations have been in place since 2012, and were last updated in November 2023.

The aim of the AML regimes globally is also influenced by the FATF. As mentioned in the introduction of the FATF 2012 Recommendations, FATF's ultimate aim is to safeguard and protect 'the integrity of the international financial system'. The aims of the EU AML regime are similar. For instance, recital 50 of the Fifth Anti-Money Laundering Directive (AMLD5)[24] states that the objective of the Fourth Anti-Money Laundering Directive (AMLD4)[25] is 'the protection of the financial system by means of prevention, detection and investigation of money laundering and terrorist financing'. Meanwhile, in the US, a broader purpose has been attributed to the AML rules; the detection and reporting of suspicious activity 'including the predicate offenses to money laundering and terrorist financing, such as securities fraud and market manipulation'.[26] The essence of the aim of the regime has remained the same, although the US has expressly mentioned terrorist financing in the domestic definition. At the same time, the AML/CFT field contains a multitude of stakeholders on a global scale, such as relevant regulators,[27] and other

22 FATF (2012-2023), International Standards on Combating Money Laundering and the Financing of Terrorism & Proliferation, <https://www.fatf-gafi.org/en/publications/Fatfrecommendations/Fatf-recommendations.html>.

23 Directive 2005/60/EC of the European Parliament and of the Council of 26 October 2005 on the prevention of the use of the financial system for the purpose of money laundering and terrorist financing (Text with EEA relevance), OJ L 309, 25.11.2005, p. 15–36.

24 Directive (EU) 2018/843 of the European Parliament and of the Council of 30 May 2018 amending Directive (EU) 2015/849 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing, and amending Directives 2009/138/EC and 2013/36/EU PE/72/2017/REV/1, *OJ L 156, 19.6.2018, p. 43–74.*

25 Directive (EU) 2015/849 of the European Parliament and of the Council of 20 May 2015 on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing, amending Regulation (EU) No 648/2012 of the European Parliament and of the Council, and repealing Directive 2005/60/EC of the European Parliament and of the Council and Commission Directive 2006/70/EC, OJ L 141, 5.6.2015, p. 73–117.

26 Financial Industry Regulatory Authority (FINRA), Rules & Guidance, Key Topics, Anti-Money Laundering (AML), <https://www.finra.org/rules-guidance/key-topics/aml>.

27 Government agencies responsible for creating and enforcing AML/CFT laws and regulations.

obliged entities. Obliged entities include credit institutions, financial institutions and other individual or legal entities that conduct certain business activities.[28] These entities are required by law to comply with AML/CFT rules and take relevant preventive measures. They play a very important role in AML/CFT schemes, conducting due diligence,[29] reporting suspicious transactions and keeping relevant records that must be reported to authorities. On the other hand, LEAs are also considered actors of the AML/CFT ecosystem, being responsible for investigating and prosecuting money laundering cases. For a more effective enforcement of the cases, LEAs also collaborate with obliged entities on an international level to combat cross-border money laundering activities. Finally, individuals play a dual role in this scheme. They are data subjects by nature,[30] while also being customers of financial institutions and other business entities subject to AML/CFT obligations, and they are required to provide their personal data and information about their transactions.[31]

In terms of national AML/CFT policies, Financial Intelligence Units (FIUs) play an important role. FIUs are central, national agencies responsible for receiving, analysing and disseminating financial information about suspected activities of financial crime and terrorism financing to the competent authorities in order to combat money laundering and terrorism financing.[32] FIUs are the central national point for receiving and analysing suspicious transaction reports and other relevant financial information from obliged entities. FIUs can also use their legal powers to request information from other competent authorities, government agencies or third parties, if relevant to their investigations or analysis. FIUs can further request information from foreign counterparts through mutual legal assistance agreements or other forms of international cooperation, such as the Egmont Group,[33] a global network of FIUs that facilitates the exchange of financial intelligence information.

28 As set forth in Article 2(1)(3) AMLD 4.

29 Identification and verification of the customers' identity, and, when the customers are legal entities, the beneficial owners of the customers. This procedure includes obtaining and verifying customer information, such as name, address, and identification documents, and assessing the risk associated with the customer's activities to detect and prevent potential money laundering or terrorist financing.

30 Middleton, A. M. (2010). Data-intensive technologies for cloud computing. *Handbook of cloud computing*, 83-136.

31 De Koker, L. (2006). Money laundering control and suppression of financing of terrorism: some thoughts on the impact of customer due diligence measures on financial exclusion. *Journal of financial crime*, *13* (1), 26-50.

32 European Commission, Directorate-General for Migration and Home Affairs, *Fight against the financing of terrorism*, <https://home-affairs.ec.europa.eu/policies/internal-security/counter-terrorism-and-radicalisation/fight-against-financing-terrorism_en>.

33 The Egmont Group of Financial Intelligence Units is an international organisation that facilitates

The intricacy created by the different actors and the variety of laws creates a complicated AML/CFT ecosystem. In light of this complex environment and aiming to clarify the existing situation and enhance the European anti-money laundering rules, the European Commission presented a new package of legislative proposals in July 2021.[34] Besides the general rules on the processing and the cross-border transferring of data, established in the General Data Protection Regulation (GDPR), specific issues relating to personal data movement within the EU are regulated through these new AML/CFT legislation proposals, where tailor-made data protection provisions can be found.[35] The new legislative package includes proposals for: (a) a regulation on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing (AML Regulation), (b) a sixth AML Directive (AMLD6), (c) a regulation establishing the Authority for Anti-Money Laundering and Countering the Financing of Terrorism (AMLA Regulation), and (d) a regulation on information accompanying transfers of funds and certain crypto assets. This package aims to achieve direct applicability of an AML regulatory framework that is directly applicable to the national jurisdictions of Member States (through regulations), and EU-level supervision with an EU-wide AML/CFT supervisory system.[36] However, compliance of the new package with the regulatory frameworks that cover the protection and transfers of personal data is still unclear.[37]

cooperation and intelligence sharing between national financial intelligence units to investigate and prevent money laundering and terrorist financing. <https://egmontgroup.org/>.

34 European Commission, Directorate-General for Financial Stability, Financial Services and Capital Markets Union, *Anti-money laundering and countering the financing of terrorism legislative package*, <https://ec.europa.eu/info/publications/210720-anti-money-laundering-countering-financing-terrorism_en>, (20 July 2021).

35 See, for example, Chapter VI (*Data Protection and Record-retention*) of the Proposal for a Regulation of the European Parliament and of the Council on the prevention of the use of the financial system for the purposes of money laundering or terrorist financing.

36 European Commission, Directorate-General for Financial Stability, Financial Services and Capital Markets Union, *Anti-money laundering and countering the financing of terrorism legislative package*, <https://ec.europa.eu/info/publications/210720-anti-money-laundering-countering-financing-terrorism_en>, (20 July 2021).

37 Communication from the Commission on an Action Plan for a comprehensive Union policy on preventing money laundering and terrorist financing (2020/C 164/06).

## 3. International Transfers of Data in the Context of AML/CFT between the EU and the US: A Thorny Issue

International personal data transfers can be conducted under several circumstances and for numerous purposes. In the context of AML/CFT activities, transfers of personal data are essential for reporting suspicious transactions to authorities in different jurisdictions. Personal data transfers can be crucial for AML/CFT purposes within the context of collaboration between AML/CFT actors, as they enable the detection of suspicious activities, compliance with regulatory requirements, enhanced due diligence, and the investigation and prosecution of money laundering and other financial crimes.[38] Experts hold diametrically opposed opinions regarding transfers of personal data for AML/CFT purposes, as some believe that personal data should be transferred much more flexibly on a worldwide basis, especially for law enforcement matters.[39] Others place much greater importance on the fundamental rights of privacy and data protection.[40]

The fundamental right of data protection is highly regarded in European AML/CFT legislation. More specifically, in the current applicable AML directive, the AMLD5, there is clear reference to the data protection laws pertinent to processing data for AML purposes. These are the General Data Protection Regulation (GDPR)[41] and the Law Enforcement Directive (LED).[42]

In terms of personal data transfers, there is another level of collaboration between the EU and the US, since they maintain a Mutual Legal Assistance Treaty (MLAT).[43,44] Mutual legal assistance is a form of cooperation between countries aiming to collect and

38 European Commission, EU context of anti-money laundering and countering the financing of terrorism, <https://finance.ec.europa.eu/financial-crime/eu-context-anti-money-laundering-and-countering-financing-terrorism_en>, 22 February 2022.

39 Maguire, M. (2000). Policing by risks and targets: Some dimensions and implications of intelligence-led crime control. *Policing and Society: An International Journal*, *9* (4), 315-336.

40 Enshrined in Articles 7 and 8 of the European Charter of Fundamental Rights, respectively.

41 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) (OJ L 119, 4.5.2016, p. 1).

42 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 (OJ L 119, 4.5.2016, p. 89).

43 U.S. Department of Justice, Criminal Division, Office of International Affairs, *Mutual Legal Assistance Treaties Of The United States* (April 2022).

44 Agreement on mutual legal assistance between the European Union and the United States of America, OJ L 181, 19.7.2003, p. 34–42.

exchange information, and request or provide evidence located in one country to assist in criminal investigations or proceedings in another.[45] However, MLATs are largely considered bureaucratic, slow and ineffective for exchanging data.[46] Concerning the future AML/CFT legislative proposals, the European Data Protection Board (EDPB)[47] has encouraged the European Commission to propose specific provisions to adapt data protection rules more smoothly. The EDPB considers that in case the AML/CFT laws are not designed in a balanced and proportionate manner, respecting the fundamental right to data protection, legal uncertainties will persist and the AML-CFT framework would be vulnerable.[48]

In view of the above, it is questionable whether the new legislative package will be able to ensure that money laundering is dealt with effectively, while balancing law enforcement goals and the fundamental right of personal data protection. Subsequently, an arising question is how relevant actors will be practically affected by the new regime. This derives from the fact that the new legislative package does not clearly reflect on either data sharing or transborder personal data transfers for AML/CFT purposes;[49] the new EU AML law follows the current legislative path. As indicated in recital 85 of the AMLD 6 draft, the processing (including transfers) of data will rely on the GDPR and the LED. Considering that the unclear situation has persisted until now even with the current legislative instruments deployed, the new legislative package might have missed the opportunity to address this matter properly.

Under the present circumstances, there seems to be a significant gap in the current regime. It ought to be covered before actual risks are created concerning the accountability of all stakeholders involved and the fundamental rights of privacy and data protection. This gap is the absence of a fundamental rights-respecting, yet effective regulatory environment for governing transatlantic personal data transfers in a way that achieves a fair balance between fundamental rights and law enforcement efficiency in the AML/CFT field.

---

45 European Commission, *Mutual legal assistance and extradition*, <https://commission.europa.eu/law/cross-border-cases/judicial-cooperation/types-judicial-cooperation/mutual-legal-assistance-and-extradition_en>.

46 Hill, J. F. (2015). Problematic alternatives: MLAT reform for the digital age. *Harvard Law School: National Security Journal*, 1.

47 The European Data Protection Board is a European Union independent body with juridical personality whose purpose is to ensure consistent application of the GDPR and to promote cooperation among the EU's data protection authorities.

48 European Data Protection Board, *EDPB letter to the European Commission on the protection of personal data in the AML-CFT legislative proposals* (19 May 2021)

49 European Data Protection Board, *EDPB letter to the European Parliament, the Council, and the European Commission on data sharing for AML/CFT purposes in light of the Council's mandate for negotiations* (28 March 2023)

## 4. Open Issues in the Existing Frameworks; Differences in the AML/CFT and Data Protection Regimes

Concerning transatlantic personal data transfers between the EU and the US for AML/CFT purposes, the existing regime seems to be problematic for a series of reasons related to both data protection and the AML/CFT fields.

From a data protection point of view, the respective legislative frameworks in the two jurisdictions maintain different approaches. The EU has established a comprehensive data protection legal framework that consists of several Directives and Regulations applicable to all EU Member States. The most important of these is the GDPR. In contrast, the US maintains a patchwork of federal and state laws that govern data protection. Although there are important federal data protection laws,[50] there is no single comprehensive federal data protection law that applies to all sectors and industries. This affects the actualisation of data transfers.

Between 2016 and 2020, personal data transfers between the EU and the US were conducted based on the Privacy Shield. This was a framework for data transfers between the two jurisdictions. However, in July 2020, the EU-US Privacy Shield framework was invalidated by the *Schrems II* judgment of the Court of Justice of the European Union (CJEU).[51] The CJEU ruled that the Privacy Shield framework did not provide sufficient protection for European citizens' personal data when transferred to the US. It ruled that it did not adequately address US surveillance practices and the lack of effective judicial remedies for European individuals whose data was accessed by U.S. authorities.[52] This invalidation made transfers of personal data between the EU and the US more complicated.

However, this is not the only framework under which personal data are transferred. There are currently multiple frameworks for data transfers for law enforcement purposes. For instance, Chapter V of the Law Enforcement Directive (LED)[53] concerns

50 Including the Gramm-Leach-Bliley Act (GLBA), which applies to financial institutions and requires them to protect the confidentiality and security of individuals' personal financial information or the Fair Credit Reporting Act (FCRA), which regulates the collection, use, and dissemination of consumer credit information by credit reporting agencies.

51 Case C-311/18, Judgment of the Court (Grand Chamber) of 16 July 2020, Data Protection Commissioner v Facebook Ireland Limited and Maximillian Schrems, Request for a preliminary ruling from the High Court (Ireland).

52 Tracol, X. (2020). "Schrems II": The return of the privacy shield. *Computer Law & Security Review*, *39*, 105484.

53 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, OJ L 119, 4.5.2016, p. 89–131.

transfers of personal data to third countries for law enforcement purposes by competent authorities (as defined in the LED). Meanwhile, the European regulation regarding electronic evidence (e-evidence) aims to facilitate cross-border access to electronic evidence in criminal investigations between the EU and third countries, including the US,[54] and is expected to play a large role in data transferring for law enforcement purposes. On the other side of the Atlantic, the US enacted the CLOUD Act.[55] This is a federal law designed to address issues related to law enforcement access to electronic data processed by US-based technology companies that are stored overseas, as well as data processed by foreign tech companies accessible in the US. The above are subject to the EU-US Data Protection Umbrella Agreement. This is a framework for the protection of personal data exchanged for the purpose of preventing, investigating, detecting or prosecuting criminal offenses, including terrorism. It provides a comprehensive set of data protection safeguards for the exchange of personal data between the EU and the US in the law enforcement context.[56] All these frameworks exist simultaneously and intercorrelate between themselves and the GDPR.

Apart from the above data protection-related matters, there are also distinct differences in the AML/CFT laws between the two jurisdictions. At first glance, both the EU and the US have adopted a risk-based approach to AML/CFT, meaning that they require financial institutions and other obliged entities to identify and assess the money laundering risks they face, and take measures to mitigate those risks. This risk-based approach stems from the FATF recommendations. These state that this approach enables countries to adopt a more flexible set of measures to more effectively target their resources and implement commensurate preventive measures to focus their efforts.[57] However, the EU approach is more prescriptive. It is set out in multiple directives and regulations. These are vertically applicable to all obliged entities, with specific requirements for customer due diligence, beneficial ownership and reporting of suspicious transactions. Comparatively, the US allows more flexibility for institutions to tailor their AML/CFT programs to their specific risk profiles.[58]

54 Regulation (EU) 2023/1543 of the European Parliament and of the Council of 12 July 2023 on European Production Orders and European Preservation Orders for electronic evidence in criminal proceedings and for the execution of custodial sentences following criminal proceedings, PE/4/2023/REV/1, OJ L 191, 28.7.2023, p. 118–180.

55 The Clarifying Lawful Overseas Use of Data Act (H.R. 4943).

56 Council Decision (EU) 2016/920 of 20 May 2016 on the signing, on behalf of the European Union, of the Agreement between the United States of America and the European Union on the protection of personal information relating to the prevention, investigation, detection, and prosecution of criminal offences, OJ L 154, 11.6.2016, p. 1–2.

57 Recommendation No. 1, FATF Recommendations.

58 Financial Crimes Enforcement Network U.S. Department of the Treasury, *Anti-Money Laundering and Countering the Financing of Terrorism National Priorities* (30 June 2021)

Although the EU and the US have similar goals and principles for AML/CFT, their legal frameworks and mechanisms differ in significant ways. This can create a number of problems related to personal data protection, which will affect all abovementioned stakeholders. As the legal framework for data transfers for AML/CFT purposes is complex, this creates uncertainty and confusion for legal entities, organisations and individuals. Since the legal frameworks for data transfers vary across jurisdictions, this can create challenges for multinational companies that operate in several countries as well as for LEAs collaborating for AML/CFT matters. For example, the definition of a 'suspicious transaction' may vary between jurisdictions. This can make it difficult to apply consistent AML/CFT policies and procedures across a global organisation.[59] Such procedures would include transfers of data for AML/CFT purposes. Furthermore, from a business perspective, compliance burdens are already high, as the GDPR rules on data transfers require companies to implement appropriate safeguards to protect personal data. These are costly and time-consuming.[60] Moreover, transfers of personal data for AML/CFT purposes may create risks to privacy and security, particularly if there is no adequate protection.[61] Data breaches or unauthorised disclosures of personal data can result in identity theft, financial fraud and other forms of harm.[62] In conclusion, personal data transfers for AML/CFT purposes may create numerous challenges related to data protection and privacy. It is unclear whether there is a way to deal with these challenges, and this will most likely be a discussion in years to come. The legal landscape may need revision, calling for specialised rules for data transfers for said purposes, since there is an uncertainty as to whether the current, general ones are fit for their purpose, in terms of both efficiency and with respect to fundamental rights.

59 Europol, Financial Intelligence Group, *From Suspicion to Action | Converting financial intelligence into greater operational impact* (2017), <https://www.europol.europa.eu/sites/default/files/documents/ql-01-17-932-en-c_pf_final.pdf>.

60 Digital Europe, *Data transfers in the data strategy: Understanding myth and reality* (16 June 2022) <https://www.digitaleurope.org/resources/data-transfers-in-the-data-strategy-understanding-myth-and-reality/>.

61 European Commission, Directorate-General for Financial Stability, Financial Services and Capital Markets Union, *Anti-money laundering and countering the financing of terrorism legislative package*, <https://ec.europa.eu/info/publications/210720-anti-money-laundering-countering-financing-terrorism_en>, (20 July 2021).

62 Peretti, K. K. (2008). Data breaches: What the underground world of "carding" reveals. *Santa Clara Computer & High Tech. LJ*, *25*, 375.

## 5. Conclusions: Need for Elucidation of the Legal Regime Regarding International Data Transfers Between the EU and the US in the Field of AML/CFT

This chapter has been dedicated to the exploration of the current situation of data transfers between obliged entities, LEAs and FIUs in a transborder context. It aimed to highlight the urgent need to clarify the existing regime on transfers for AML/CFT purposes. The current data protection regime does not comprehensively govern transfers of personal data between actors of the AML/CFT field. There is a pressing need to end the uncertainty around the rights of those whose data are processed and the obligations of actors involved in the processing. Hence, there is a gap in the making of the new regime. It must be covered as soon as possible as it bears risks on the accountability of all stakeholders involved. This gap is the absence of a concrete regulatory environment for governing data sharing in the AML/CFT field. More specifically, this refers to the absence in collaboration schemes, where FIUs, banks, law enforcement authorities (LEAs) and other entities work together. The regime may need to be enhanced, taking two main matters into consideration: respect for the fundamental rights of privacy and data protection, and effective law enforcement AML/CFT mechanisms.

The principal goal of future research should be to produce a new holistic approach to the data transfers for AML/CFT purposes between the EU and the US. This research aims to accurately indicate and demarcate as many aspects of the existing problems as possible, as well as to provide regulatory recommendations to solve these. However, future research has the potential to be ground-breaking and to provide a generalised global approach on data transfers for financial crime – and more specifically for AML/CFT-purposes between different jurisdictions.

CHAPTER

# IV

# Investigative Genetic Genealogy: An Emerging Legal Concern in Europe?

**Taner Kuru**[1]

https://doi.org/10.26116/6gde-4515

1 PhD Researcher at Tilburg Institute for Law, Technology, and Society (TILT), Tilburg Law School, Tilburg University.

## 1. Introduction

DNA was first used as evidence in a criminal case in 1986 in the United Kingdom.[2] In this case, it did not only help investigators to identify the killer, but it also led an innocent suspect to be freed after spending more than three months in custody.[3] Since then, law enforcement authorities (hereinafter 'LEA') around the world have used this powerful source of evidence to solve their cases. To be able to benefit from the promises of DNA in criminal justice, they also started establishing national forensic DNA databases in which they store DNA profiles extracted from the genetic material of particular groups of people, such as suspects, arrestees, convicted individuals, and missing people.[4]

In Europe, national legislatures provide specific laws that regulate DNA retention and national forensic DNA databases. However, due to their intrusive nature regarding the rights and freedoms of individuals concerned, especially regarding the right to privacy, these laws have been subjected to judicial scrutiny in many instances. The European Court of Human Rights (hereinafter 'the ECtHR' or 'the Strasbourg court') has examined these provisions from various jurisdictions in the last two decades. Very recently, the Court of Justice of the European Union (hereinafter 'the CJEU' or 'the Luxembourg court') interpreted the provisions in the EU legal framework related to DNA retention for the first time, and it reached conclusions that follow the same line taken by the settled case law of the ECtHR.[5] As a result, legal restraints attached to DNA retention and the governance of national forensic DNA databases in Europe have been shaped by the jurisprudence of these two European courts.

However, a recent ground-breaking advancement in criminal investigations may challenge this set of legal requirements. In 2018, the investigators of the Golden State Killer case in the United States announced that they identified the serial killer whom they had been trying to catch for more than forty years thanks to a novel investigation technique.[6] Although the investigators of this case had the crime scene DNA, they could not find a match through CODIS,[7] which could have allowed them to identify the Golden

2 Cobain, I. (2016). Killer breakthrough: the day DNA evidence first nailed a murderer. *The Guardian*, 7.

3 Ibid.

4 Santos, F., Machado, H., & Silva, S. (2013). Forensic DNA databases in European countries: is size linked to performance?. *Life Sciences, Society and Policy*, *9*, 1-13. Amankwaa, A. O. (2018). Forensic DNA retention: public perspective studies in the United Kingdom and around the world. Science & Justice, 58(6), 455-464. Butler, J. M. (2023). Recent advances in forensic biology and forensic DNA typing: INTERPOL review 2019–2022. *Forensic Science International: Synergy*, *6*, 100311.

5 Case C-205/21, V.S. v. Ministerstvo na vatreshnite raboti, Glavna direktsia za borba s organiziranata prestapnost, 26.6.2023, ECLI:EU:C:2023:49.

6 Lussenhop, J. (2018). Golden state killer: the end of a 40-year hunt?. *BBC News*, *29*. Arango, T., Goldman, A., & Fuller, T. (2018). To catch a killer: A fake profile on a DNA site and a pristine sample. *The New York Times*, *27*.

7 The Combined DNA Index System (CODIS) is the software used to, among others, manage the

State Killer.[8] This was because neither the suspect nor his biological family members' DNA were in CODIS.[9] Therefore, the investigators of this case decided to try something different: they uploaded the crime scene DNA to GEDmatch, an online genetic genealogy platform where users upload their genetic test results provided by direct-to-consumer genetic testing services to find their biological relatives,[10] which led to the identification of Joseph James DeAngelo as the Golden State Killer.[11]

After this story made its way to international news headlines, investigators in the United States and several other countries started using this technique, called investigative genetic genealogy (hereinafter 'IGG'), in their investigations.[12] The Swedish LEA became the first in Europe to experiment with IGG in 2019. This led to the second biggest case in the country's history being closed after remaining unsolved for over fifteen years.[13] Recently, the Dutch LEA also announced its interest in this technique and has already been green-lighted by the court for two pilot cases.[14] Furthermore, although no IGG cases have been reported so far, the Danish parliament passed a law that allows Danish LEA to benefit from this technique for cases related to certain crimes (e.g. murder and sexual assault) when several conditions are met.[15]

DNA data stored in the National DNA Index System (NDIS), the national DNA forensic database of the United States of America, that is maintained by the Federal Bureau of Investigation (FBI). See *Frequently Asked Questions on CODIS and NDIS*, FBI, <https://www.fbi.gov/how-we-can-help-you/dna-fingerprint-act-of-2005-expungement-policy/codis-and-ndis-fact-sheet>.

8 *Privacy concerns after public genealogy database used to ID "Golden State Killer" suspect*, CBS News (27 April 2018), <https://www.cbsnews.com/news/privacy-concerns-after-public-genealogy-database-used-to-id-golden-state-killer-suspect/>; Ali, K. (2022) *The Science Behind the Golden State Killer's Capture*, Medium (8 June 2022), <https://medium.com/@kasifaliwdr/the-science-behind-the-golden-state-killers-capture-ffac9e61bb6d>.

9 Lussenhop, *supra* n. 6.

10 *About GEDmatch*, GEDmatch, <https://www.gedmatch.com/about/>.

11 Selk, A. (2018). The ingenious and 'dystopian'DNA technique police used to hunt the 'Golden State Killer'suspect. *Washington Post*, *28*.

12 Granja, R. (2023). Citizen science at the roots and as the future of forensic genetic genealogy. *International Journal of Police Science & Management*, *25*(3), 250-261.

13 Tillmar, A., Fagerholm, S. A., Staaf, J., Sjölund, P., & Ansell, R. (2021). Getting the conclusive lead with investigative genetic genealogy–A successful case study of a 16 year old double murder in Sweden. *Forensic science international: genetics*, *53*, 102525.

14 Endedijk, B. & Van den Berg, E. (2023). *OM wil particuliere dnadatabanken uit de VS inzetten bij het oplossen van cold cases*, NRC (5 March 2023), <https://www.nrc.nl/nieuws/2023/03/05/om-wil-particuliere-dna-databanken-uit-de-vs-inzetten-bij-het-oplossen-van-cold-cases-a4158669>; Endedijk, B. & Van den Berg, E. (2023). *Justitie gaat commerciële dnadatabases gebruiken om twee cold cases alsnog op te lossen*, NRC (28 September 2023) <https://www.nrc.nl/nieuws/2023/09/28/justitie-gaat-commerciele-dna-databases-gebruiken-om-twee-cold-cases-alsnog-op-te-lossen-a4175697>.

15 *B 15 Forslag til folketingsbeslutning om, at dansk politi skal kunne bruge genetisk slægtsforskning i efterforskning af drab og grov personfarlig kriminalitet (borgerforslag).*, Folketinget, <https://www.ft.dk/samling/20222/beslutningsforslag/b15/index.htm>.

Given this growing interest, it is evident that the number of cases in which IGG is used will increase in Europe in the coming years.[16] Accordingly, this chapter provides an overview of the current legal framework regarding DNA retention and national forensic DNA databases in Europe and assesses how European LEAs' emerging interest in IGG might challenge this framework. In this regard, it first presents the legal requirements for DNA retention and national forensic DNA databases in Europe set forth by the jurisprudence of the ECtHR and the CJEU. It then introduces IGG and highlights that, despite its success and promises, it raises novel ethical and legal challenges, especially by allowing LEAs to practically circumvent the legal restraints for DNA retention set forth by the two European courts. Therefore, the chapter concludes by identifying IGG as an emerging legal concern in Europe and calls for urgent action on this matter.

## 2. Legal Restraints Attached to DNA Retention and National Forensic DNA Databases in Europe

### 2.1 An Overview of the ECtHR Jurisprudence

DNA is considered crucial evidence in criminal investigations, and its collection and use in criminal proceedings are regulated by national legislation in Europe. These rules, however, have been subjected to judicial scrutiny due to the intrusive nature of these practices regarding the rights and freedoms of the individuals who are subjected to these practices. While the ECtHR jurisprudence provides more extensive content on this matter, given that it has been deciding on relevant cases over the last two decades already, the CJEU also had the opportunity to provide its view on the matter very recently. Accordingly, it is vital to analyse these decisions to understand the legal restraints attached to DNA retention and national forensic DNA databases in Europe.

16 It is reported that the Norwegian police also benefited from genealogy databases in the Knut Kristiansen case. However, from the available sources, it unclear whether the databases used in this case were those of genetic genealogy databases, such as GEDmatch or FamilyTreeDNA, since the sources do not mention them but Oslo University Hospital and National Archives. See, *Oslo Police Use Genetic Genealogy for the First Time to ID Murder Suspect*, Forensic (8 February 2023) <https://www.forensicmag.com/594329-Oslo-Police-Use-Genetic-Genealogy-for-the-First-Time-to-ID-Murder-Suspect/>. Matre, J. Solheim, E.K., Fausko, L. Eggen, S. & Åsgard, A. F. M. *Knut Kristiansen funnet drept i 1999 – mener de vet hvem drapsmannen er*, VG (2 February 2023, 19:04 CET) https://www.vg.no/nyheter/innenriks/i/q1kgM1/innkaller-til-pressekonferanse-om-ny-utvikling-i-eldre-drapssak. Furthermore, the Biometrics and Forensics Ethics Group of the UK's Home Office released a feasibility report on the use of IGG in the UK in 2020; however, no further action or cases have been reported so far from the UK on this matter. See, *Should we be making use of genetic genealogy to assist in solving crime?, A report on the feasibility of such methods in the UK*, The Biometrics and Forensics Ethics Group (September 2020).

In its jurisprudence, the ECtHR has repeatedly acknowledged that DNA evidence is crucial for the criminal justice system and for LEAs.[17] Nevertheless, it underlined that the retention of DNA samples constitutes an intrusion on the individuals' right to privacy,[18] given that genetic information contains very sensitive information about them and considering the unforeseen future uses and abuses of this information.[19] Lastly, the Strasbourg court highlighted that the fact that the extracted DNA profiles could be used to identify the genetic relationship between individuals should be enough to conclude that DNA retention infringes the right to privacy regardless of the frequency of familial searches, implemented safeguards, and the likelihood of detriment.[20] Consequently, the ECtHR underlined that such interferences could be justified as long as they are in accordance with the law, pursue a legitimate aim, and are necessary in a democratic society to achieve the aims concerned as per paragraph 2 of Article 8 of the European Convention on Human Rights (hereinafter "the ECHR").[21]

First, it is observed in these cases that the ECtHR considered the 'in accordance with law' criterion very closely related to the 'necessary in a democratic society' criterion, so it did not examine the 'quality of law' threshold,[22] unless the practice clearly missed a specific legal basis in national law.[23] Second, it repeatedly acknowledged that LEAs retaining DNA data pursues a legitimate aim because such data hold crucial importance for detecting and preventing crime and, as such, protecting the rights and freedoms of others.[24] More interestingly, the Strasbourg court considered DNA retention as pursuing a legitimate aim since it serves a broader purpose of assisting in the identification of future offenders or solving cold cases, which are unsolved criminal cases awaiting new evidence.[25] Accordingly, the ECtHR opined that society's interests in crime prevention

17 S. and Marper v. the United Kingdom [GC], no. 30562/04 and 30566/04, §105; Peruzzo and Martens v. Germany, no 7841/08 and 57900/12, §42; Trajkovski and Chipovski v. North Macedonia, no. 53205/13 and 63320/13, §51.

18 van der Velden v. the Netherlands, no. 29514/05; S. and Marper, *supra* n. 17, §71-77; W. v. the Netherlands, no. 20689/08; Aycaguer v. France, no. 8806/12, §33; Gaughran v. the United Kingdom, no. 45245/15, §63; Trajkovski and Chipovski, *supra* n. 17, §43; Dragan Petrovic v. Serbia, no. 75229/10, §79.

19 van der Velden, *supra* n. 18; S. and Marper, *supra* n. 17, §71-73; Aycaguer, *supra* n. 18, §33.

20 S. and Marper, *supra* n. 17, §75; Gaughran, *supra* n. 18, §81; *See also* Peruzzo and Martens, *supra* n. 17, §42.

21 W. v. the Netherlands, *supra* n. 18; Peruzzo and Martens, *supra* n. 17, §34; Gaughran, *supra* n. 18, §71.

22 S. and Marper, *supra* n. 17, §99; Peruzzo and Martens, *supra* n. 17, §39; Gaughran, *supra* n. 18, §73; Trajkovski and Chipovski, *supra* n. 17, §48.

23 Dragan Petrovic, *supra* n. 18, §81-84.

24 van der Velden, *supra* n. 18; S. and Marper, *supra* n. 17, §100; W. v. the Netherlands, *supra* n. 18; Peruzzo and Martens, *supra* n. 17, §40; Aycaguer, *supra* n. 18, §36; Gaughran, *supra* n. 18, §75; Trajkovski and Chipovski, *supra* n. 17, §49.

25 S. and Marper, *supra* n. 17, §100; Peruzzo and Martens, *supra* n. 17, §40; Gaughran, *supra* n. 18, §75-93; Trajkovski and Chipovski, *supra* n. 17, §49.

could outweigh the interests of individuals who are subjected to these practices.[26] However, at this point, the ECtHR has referred to the particular sensitivity of genetic information and emphasised that the use of modern scientific techniques in criminal justice systems and the investigation of cold cases should not lead to diminishing the right to privacy of the individuals concerned. Instead, these techniques should be implemented after conducting a carefully considered balancing test.[27]

Hence, the Strasbourg court required these practices to be accompanied with certain safeguards, such as:[28]

- respecting the data minimisation principle,
- having specific purposes for the retention,
- maintaining limited storage periods,
- drawing up clearly defined governance schemes that include minimum safeguards regarding third-party access,
- preserving data integrity and confidentiality,
- maintaining deletion and destruction procedures.

Accordingly, when national legislation fell short of satisfying these criteria, the ECtHR decided that DNA retention violated the right to privacy of the individuals concerned. While reaching this conclusion, the ECtHR considered various elements, such as:[29]

- the existence of a blanket and indiscriminate DNA retention practice,
- that DNA samples could be taken from individuals, irrespective of the nature or gravity of their offences,
- that the age of the individuals from whom the DNA samples could be taken was not considered,
- the retention of data for a practically unlimited period,
- the failure to give individuals concerned an opportunity to have their data removed,
- the absence of an independent review for the retention.

When, however:

- the national legislation only allowed retention from individuals convicted of an offence with a certain gravity,
- the retention period was limited to a reasonable timeframe,
- an adequate data governance scheme was drawn, or

26 S. and Marper, *supra* n. 17, §104.
27 S. and Marper, *supra* n. 17, §112; Gaughran, *supra* n. 18, §93.
28 S. and Marper, *supra* n. 17, §99-107; Aycaguer, *supra* n. 18, §38.
29 S. and Marper, *supra* n. 17, §119-125; Aycaguer, *supra* n. 18, §42-45; Gaughran, *supra* n. 18, §81-96; Trajkovski and Chipovski, *supra* n. 17, §52-54.

- the DNA sample had to be destroyed once it completed its purpose, and the individuals concerned had the opportunity to ask for their data to be deleted,

the ECtHR considered that there was no blanket and indiscriminate DNA retention, and adequate safeguards were ensured for effective protection against the misuse and abuse of the retained data, even when the individual concerned was a minor.[30] In conclusion, it can be said that the ECtHR did not consider DNA retention as such against the ECHR in its jurisprudence. However, it stressed the importance of implementing sufficient legal and practical safeguards against the misuse and abuse of the retained data.

### 2.2 DNA Retention by EU LEAs: What Does the LED Say?

In the EU legal framework, the Law Enforcement Directive (hereinafter 'LED')[31] regulates the processing of personal data by LEAs.[32] The LED lists genetic data as one of the special categories of personal data in its Article 10, therefore subjecting its processing to a stricter protection regime. According to this provision, processing of genetic data is allowed only where strictly necessary, subject to appropriate safeguards for the rights and freedoms of the data subject; where authorised by Union or Member State law, to protect the vital interests of the data subject or of another natural person; or where the processing relates to data which are manifestly made public by the data subject. However, the LED falls short of explaining further what should be understood from the 'strictly necessary' and 'appropriate safeguards' conditions mentioned thereof. Nevertheless, the CJEU had the chance to interpret some of the conditions listed in Article 10 LED with the *V.S.* case in the DNA retention context.[33]

V.S. was accused of participating in a criminal organisation related to tax fraud, an intentional criminal offence subject to public prosecution under Bulgarian law.[34] According to Bulgarian law, the LEA is obliged to create police records of individuals who are accused of an intentional criminal offence subject to public prosecution, which contain, among others, DNA profiles extracted from the samples collected from them.[35] If the individuals concerned oppose the collection of their samples to create their DNA profiles,

30 W. v. the Netherlands, *supra* n. 18; Peruzzo and Martens, *supra* n. 17, §44-47.

31 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, OJ L 119, 4.5.2016.

32 Art. 1(1) LED

33 V.S., *supra* n. 5.

34 V.S., *supra* n. 5, §36.

35 V.S., *supra* n. 5, §31.

they are subjected to a compulsory collection after authorisation from the criminal court.[36] During the criminal proceedings against her, V.S. refused to cooperate with the LEA to provide her sample.[37] Following this refusal, Bulgarian LEA asked the criminal court to authorise them to collect her sample to create her DNA profile as part of her police record.[38] However, the criminal court requested a preliminary ruling from the CJEU regarding the compatibility of the relevant provisions of the Bulgarian legislation allowing such compulsory collection with the LED and the Charter of Fundamental Rights of the European Union (hereinafter 'the Charter').[39]

In its decision, the CJEU stated that the collection of DNA samples in criminal proceedings serves the aims of prevention, investigation, detection, and prosecution of criminal offences, which are recognised as objectives of general interest in the EU legal framework.[40] With this background, it is acknowledged that Member State legislation providing for the compulsory collection of genetic data of individuals to create their DNA profiles is compatible with the LED when sufficient evidence is collected that the individual concerned is guilty of an intentional offence.[41] It was nevertheless underlined that Member State legislation should provide effective judicial remedies to challenge the compulsory collection of genetic data.[42] More importantly, the CJEU recalled Article 10 LED and stressed that the processing of genetic data by LEAs is subject to stricter protection; therefore, such processing should take place 'only where strictly necessary'.[43]

At this point, the CJEU stated that, while evaluating whether the processing is 'strictly necessary', it must be questioned whether such processing complies with data protection principles, such as the purpose limitation and data minimisation principles.[44] Accordingly, the purposes of DNA retention should be specific, explicit, and legitimate, and the collected data should be adequate, relevant, and limited to what is necessary.[45] Hence, the purposes of DNA retention cannot be defined too generally, and there should not be any less intrusive measures to reach the objectives pursued.[46] In light of this assessment, the CJEU decided that Member State legislation should not lead to a systematic collection of DNA of *anyone* accused of an intentional crime as it leads to indis-

36 V.S., *supra* n. 5, §31-34.
37 V.S., *supra* n. 5, §37.
38 V.S., *supra* n. 5, §37-38.
39 V.S., *supra* n. 5, §39.
40 V.S., *supra* n. 5, §97.
41 V.S., *supra* n. 5, §85-86.
42 V.S., *supra* n. 5, §96.
43 V.S., *supra* n. 5, §116-117.
44 V.S., *supra* n. 5, §122.
45 Ibid.
46 V.S., *supra* n. 5, §124-126.

criminate and blanket retention.[47] In other words, LEAs should process genetic data solely in a limited number of cases instead of treating it as a general practice.[48]

In this regard, the CJEU questioned whether taking the genetic samples of accused individuals would indeed be 'strictly necessary' when there are already serious grounds for believing that they have committed a criminal offence since, in such cases, sufficient evidence has already been collected.[49] Moreover, it was emphasised that several other elements should be considered to determine whether such processing is 'strictly necessary' to link different criminal procedures, such as the nature and gravity of the offences, particular circumstances of the offences, any link between these criminal procedures, and the criminal record of the individuals concerned.[50] Furthermore, the CJEU questioned whether, in the case at hand, less intrusive measures, such as processing non-sensitive personal data (e.g. civil status data), might achieve the purposes of the processing while still upholding the data minimization principle.[51] Following this, the CJEU stated that Member State legislation disregarding these points and consequently allowing systematic collection of genetic data of any person accused of an intentional offence subject to public prosecution is considered contrary to the Union law.[52]

In short, aligning with the ECtHR jurisprudence, the CJEU concluded that DNA retention as such is not contrary to the Union law, but that there should be certain limitations to such practices to strike a proper balance between competing interests. Unfortunately, in this case, the CJEU did not have the chance to analyse the characteristics of the required safeguards in the EU. However, it is safe to assume that the CJEU would set the same thresholds as the ECtHR. This is because the Luxembourg court's references to purpose limitation and data minimisation principles as well as to the importance of the availability of effective judicial remedies to challenge the (compulsory) retention seem very much aligned with the required safeguards shaped by the settled case law of the Strasbourg court in similar cases. This comes as no surprise, given that Article 52(3) of the Charter states that:

> In so far as this Charter contains rights which correspond to rights guaranteed by the Convention for the Protection of Human Rights and Fundamental Freedoms, the meaning and scope of those rights shall be the same as those laid down by the said

47 V.S., *supra* n. 5, §127-128.
48 V.S., *supra* n. 5, §118.
49 V.S., *supra* n. 5, §131.
50 V.S., *supra* n. 5, §132.
51 V.S., *supra* n. 5, §133.
52 V.S., *supra* n. 5, §135.

> Convention. This provision shall not prevent Union law providing more extensive protection.[53]

Consequently, it is safe to assume that if the CJEU were asked to shed light on the required safeguards attached to DNA retention, a picture similar to that of the ECtHR jurisprudence would be drawn.

Considering the above, EU LEAs should have a clear legal basis regarding DNA retention that does not lead to blanket and indiscriminate collection of genetic samples and that clearly prescribes the attached safeguards against data misuse and abuse. Yet a recent development in the context of using DNA evidence in criminal justice could challenge these well-established legal restraints. This is because LEAs may have recently discovered a new technique that allows them to practically circumvent these legal restraints.

## 3. Investigative Genetic Genealogy: Ground-breaking yet Controversial

After a chase of more than forty years with no results, the investigators of the Golden State Killer case decided to benefit from DNA profiles stored online in genetic genealogy databases, as they had not been able to identify the suspect via CODIS.[54] They uploaded the crime scene DNA to GEDmatch through a fake profile, which led them to find distant biological relatives of the Golden State Killer.[55] They then created family trees of these individuals and eventually identified Joseph James DeAngelo, a former police officer, as the Golden State Killer.[56] Unsurprisingly, the arrest of DeAngelo made its way to the international headlines, since it was the first high-profile case in which IGG was used.[57]

Although familial DNA searches were employed long before the Golden State Killer was caught, DeAngelo's arrest brought various ethical and legal questions concerning this practice to the attention of the global public. Until then, familial DNA searches were mainly done in the databases created and populated by law enforcement, with the DNA

53 At this point, it should be noted that Article 52(3) does not provide an obligation for the CJEU to follow the ECtHR jurisprudence in all circumstances, although the latter has a significant influence on the jurisprudence of the CJEU overall. *See further Brittain, S. (2015). The relationship between the EU Charter of Fundamental Rights and the European Convention on Human Rights: an originalist analysis. European Constitutional Law Review, 11(3), 482-511*; Bruno, G. C. (2014). *The Importance of the European Convention on Human Rights for the interpretation of the Charter of Fundamental Rights of the European Union* ch.4 (Brill Nijhoff); Lock, T. (2019). *Article 52 CFR*, (Manuel Kellerbauer, Marcus Klamert & Jonathan Tomkin eds., 2019).

54 Lussenhop, *supra* n. 6; Arango et al., *supra* n. 6.

55 Arango et al., *supra* n. 6.

56 Selk, *supra* n. 11.

57 Levenson, M. & Murphy, H. *Golden State Killer Suspect Offers to Plead Guilty*, The New York Times (29 June 2020), <https://www.nytimes.com/2020/03/04/us/golden-state-killer-trial.html>.

profiles created from the biological samples collected from people who were somehow involved in criminal investigations. As explained earlier, such practices are subject to certain legal restraints. These legal restraints practically limit the number of stored DNA samples in these databases, which means this method has a significant shortcoming. It provides little to no help if the suspects or their biological family members are not already in national forensic DNA databases. This significantly reduces what investigators can achieve with crime scene DNA. On the other hand, thanks to the boom in direct-to-consumer genetic testing services, millions of people around the world have shared their genetic material with the companies providing these services and have also uploaded their DNA profiles to genealogy websites.[58] Of course, these databases have created a valuable resource for investigators to identify suspects and human remains, and, unsurprisingly, they started benefiting from these databases, which led to the birth of IGG as a novel investigation technique.[59]

Nevertheless, certain concerns have been voiced against this expansion, especially regarding the privacy interests of the users of these databases and their biological relatives. For example, as was true in the Golden State Killer case, these users may be unaware that the information they share on these databases could be used by law enforcement and may not realise the possible implications of this function creep.[60] It is observed that in order to address this concern, investigators in the United States and beyond focus mainly on two databases, GEDmatch and FamilyTreeDNA,[61] as these allow their users to opt-in or opt-out from law enforcement access to their data for IGG searches. However, it has already been revealed that these databases may go against their promises. For example, it sparked an outcry when it was reported that GEDmatch allowed access to its database for a violent assault case, while the access had been restricted for investigations related to murder and sexual assault cases.[62] Further concerns were voiced concerning the fact that IGG allows LEAs to access the genetic data

58 Antonio Regaldo, *More than 26 million people have taken an at-home ancestry test*, MIT Technology Review (11 February 2019), <https://www.technologyreview.com/2019/02/11/103446/more-than-26-million-people-have-taken-an-at-home-ancestry-test/>.

59 The scope of this article is narrowed down to the cases in which IGG is used to identify suspects.

60 What was more interesting in the Golden State Killer case is that even Curtis Rogers, a partner of GEDmatch, acknowledged that he learned about the investigators' activities on their database through news articles and even urged people not to upload their genetic information to such databases or have them removed if they have concerns about such activities. *See* Keith Allen, Jason Hanna & Cheri Mossburg, *Police used free genealogy database to track Golden State Killer suspect, investigator says*, CNN (27 April 2018, 14:25 EDT), <https://edition.cnn.com/2018/04/26/us/golden-state-killer-dna-report/index.html>.

61 A direct-to-consumer genetic testing company. *See* <https://www.familytreedna.com/>.

62 Aldhous, P. *The Arrest Of A Teen On An Assault Charge Has Sparked New Privacy Fears About DNA Sleuthing*, BuzzFeed News (15 May 2019, 4:15), <https://www.buzzfeednews.com/article/peteraldhous/genetic-genealogy-parabon-gedmatch-assault>.

of individuals who have never been subject to criminal investigations in order to find potential matches with the suspects they chase, bearing the risk of these individuals and their biological relatives to be scrutinized unnecessarily. Indeed, it has been revealed that investigators targeted an innocent 73-year-old man in their search for the Golden State Killer, and they collected his DNA when he was lying in bed at a rehabilitation centre without the knowledge of his family.[63] Moreover, it has also been reported that an innocent individual was questioned and forced to provide investigators with his DNA sample regarding a homicide investigation after the crime scene DNA provided a close match with his father's DNA, which had been donated to a genealogy project.[64] Given these concerns, some authors claimed that IGG effectively reverses the presumption of innocence, while others underlined that it 'raises significant constitutional concerns'.[65] Lastly, while national forensic DNA databases are subject to various legal restraints, IGG is conducted on privately owned databases operating for purposes other than criminal investigations and subjected to different legal regimes, leading the legality and legitimacy of IGG to be questioned in the first place.

Despite these ethical and legal concerns, the interest of LEAs in IGG continues to grow exponentially. In fact, it is assumed that hundreds of (cold) cases have been closed after the arrest of DeAngelo, thanks to IGG.[66] However, in the regulatory environment that IGG is currently subject to, it is hard to conclude that a balance has been struck between the public's interest in detecting and preventing crime and the rights and interests of the individuals concerned.

---

63 Balsamini, D. *Cops took DNA from innocent man in nursing home in search for serial killer*, New York Post (28 April 2018 14:17 EDT) <https://nypost.com/2018/04/28/cops-took-dna-from-innocent-man-in-nursing-home-in-search-for-serial-killer/>.

64 Akpan, N. *Genetic genealogy can help solve cold cases. It can also accuse the wrong person.*, PBS News Hour (7 November 2019, 17:15 EDT) <https://www.pbs.org/newshour/science/genetic-genealogy-can-help-solve-cold-cases-it-can-also-accuse-the-wrong-person>.

65 *See for example*, Abrahamson, C. (2018). Guilt by genetic association: the Fourth Amendment and the search of private genetic databases by law enforcement. *Fordham L. Rev.*, *87*, 2539. Dery III, G. M. (2019). Can a Distant Relative Allow the Government Access to Your DNA: The Fourth Amendment Implications of Law Enforcement's Genealogical Search for the Golden State Killer and Other Genetic Genealogy Investigations. *Hastings Sci. & Tech. LJ*, *10*, 103. Ram, N. (2021). Investigative Genetic Genealogy and the Problem of Familial Forensic Identification. *Consumer Genetic Technologies: Ethical and Legal Considerations (Cohen, G., Farahany, N., Greely, H. T. & Shachar, C. eds., Cambridge Univ. Press)*; Goldstein, J. (2019). Guilty Until Proven Innocent: The Failure of DNA Evidence. *Drexel L. Rev.*, *12*, 597.

66 Glynn, C. L. (2022). Bridging disciplines to form a new one: the emergence of forensic genetic genealogy. *Genes*, *13*(8), 1381.

## 4. European LEAs Experimenting with IGG: Should We Be Worried?

It did not take long for IGG to be tested on the other side of the Atlantic. It was reported that the Swedish LEA initiated an IGG pilot project in 2019.[67] According to the evaluation report published by Swedish LEA after the completion of this project, the idea sparked after the Golden State Killer case made its way to the international news headlines.[68] This project aimed to test IGG to determine whether it has a potential for future and broader use.[69] It is claimed that ethical issues were discussed in-house during the preparations for this pilot project, and that a legal inquiry detailing the judicial framework accompanied by a data protection impact assessment was prepared.[70] The same report mentions that the Legal Affairs Department had some reservations about certain legal matters related to this pilot project.[71] Furthermore, the Swedish Data Protection Authority (hereinafter 'DPA') was not informed about the project before its initiation because the attached risks were mitigated, and they would be informed during the pilot project anyway.[72]

With this background, the pilot focused on the second biggest case in the country's history, which had remained unsolved for more than fifteen years. In this case an 8-year-old boy and a 56-year-old woman were killed, but the Swedish LEA had been unable to identify the perpetrator. They had crime scene DNA extracted from the murder weapon, but it did not show any match with samples stored in the national DNA database.[73] Hence, Swedish investigators decided to compare this crime scene DNA with the data available in the databases of GEDmatch and FamilyTreeDNA, since both companies allowed law enforcement access to their databases.[74] They found around 4,000 matches with distant relatives and after several months they eventually narrowed down their potential suspects list to two brothers.[75] When the DNA sample taken from one of the brothers provided a 100% match with the crime scene DNA, the Swedish LEA became the first in the EU to arrest an individual by using IGG.[76]

67 Aili Fagerholm, S., Tillmar, A., Staaf, J., Rying, M., & Ansell, R. (2021). Forensic DNA traces and genealogy: use of investigative genetic genealogy in criminal investigations. *The Swedish Police Authority, National Forensic Centre*; Tillmar et al. *supra* n. 13, 2.

68 Fagerholm et al., *supra* n. 67, 7.

69 Ibid.

70 Ibid.

71 Fagerholm et al., *supra* n. 67, 9.

72 Ibid.

73 Fagerholm et al., *supra* n. 67, 12.

74 Fagerholm et al., *supra* n. 67, 4-5.

75 Fagerholm et al., *supra* n. 67, 16-17.

76 *Sweden: man goes on trial for 2004 murder after DNA matched to genealogy site*, The Guardian (15 September 2020, 17:43 BST) <https://www.theguardian.com/world/2020/sep/15/man-on-trial-in-

In the aftermath of this pilot, the evaluation report further revealed that the Swedish DPA and the Swedish LEA had different interpretations of the related legal provisions.[77] First and foremost, the Swedish DPA opined that there was no legal ground for transferring the crime scene DNA to online genealogy databases located outside the EU, while the Swedish LEA disagreed with this opinion.[78] Moreover, the Swedish LEA argued that the personal data processed on the genealogy websites were made publicly available by the users, which provided legal ground for their processing as per Article 10(c) LED, while the Swedish DPA disagreed with this interpretation.[79] More strikingly, the Swedish DPA reminded the Swedish LEA that they should have requested a prior consultation before starting the pilot project and concluded by advising them not to continue using IGG, given that their actions were considered unlawful in light of the existing provisions.[80] In their evaluations, the Swedish LEA stressed the importance and urgency of law enforcement's use of IGG but agreed that a change in legislation would be 'the most constructive way forward'.[81] Accordingly, they referred to the ongoing efforts in amending the legal provisions related to the use of biometric tools by the Swedish LEA, which was expected to be finalised by early 2023 and provide the necessary legal ground for IGG to be used in Sweden.[82]

The Netherlands also announced its interest in initiating an IGG pilot project, and the Dutch LEA has recently been green-lighted to test this novel investigation technique in two cold cases.[83] The media reported that, like their Swedish colleagues, the Dutch investigators want to benefit from GEDmatch and FamilyTreeDNA databases since their users can agree that their data may be used for criminal investigations.[84] At the time of

sweden-for-double-murder-after-15-year-dna-wait>; de Groot, N. F., van Beers, B. C., & Meynen, G. (2021). Commercial DNA tests and police investigations: a broad bioethical perspective. *Journal of medical ethics*, *47*(12), 788-795.

77 Fagerholm et al., *supra* n. 67, 26-28.

78 Fagerholm et al., *supra* n. 67, 27.

79 Ibid.

80 Ibid. The Swedish DPA based this claim on Chapter 3§7 of the Swedish Criminal Data Act. In addition, it should be reminded that Article 28 LED also requires LEAs to consult the DPAs where a data protection impact assessment indicates that processing would result in high risks in the absence of measures taken to mitigate such risks or where the type of processing, especially when new technologies are used, involves high risk to the rights and freedoms of the data subjects.

81 Fagerholm et al., *supra* n. 67, 27-28.

82 Fagerholm et al., *supra* n. 67, 28. At the time of writing this chapter, the legislative procedure for the proposed law (*Biometri – för en effektivare brottsbekämpning*, Staten Offentliga Utredningar, SOU 2023:32, 2023, <https://www.regeringen.se/contentassets/08af26f9ffca495b8bd9c2f0c77e169d/biometri--for-en-effektivare-brottsbekampning-sou-2023-32.pdf>) was still ongoing.

83 Endedijk & van den Berg (September 2023), *supra* n. 14; *Police to use private DNA banks to try to solve two cold cases*, Dutch News (6 March 2023) <https://www.dutchnews.nl/2023/03/police-to-use-private-dna-banks-to-try-to-solve-two-cold-cases/>.

84 Ibid.

writing this chapter, it is yet to be seen whether a positive result is achieved in these cases, and what the legal and societal outcomes of the Dutch LEA using this technique will be. Lastly, the Danish parliament also passed a law allowing the Danish LEA to use IGG in their investigations for certain cases (e.g. murder and sexual assault), if several conditions are met.[85]

Given this growing interest, it is safe to assume that more European LEAs will benefit from privately-owned genetic genealogy databases in their efforts to identify their suspects. However, certain ethical and legal challenges must be addressed to ensure this expansion does not come at the expense of the fundamental rights and interests of the individuals concerned. For example, while Swedish and Dutch LEAs have claimed they have the legal grounds to perform IGG, the Swedish pilot case evaluation report clearly shows that the Swedish DPA disagreed with this conclusion. Furthermore, even if a clear legal ground is established for such operations, it is still questionable whether the safeguards attached to DNA retention as developed by the jurisprudence of the two European courts will be maintained, since this process is effectively outsourced. In other words, it is open to debate whether the European LEAs will practically circumvent these legal restraints on DNA retention or whether they will ensure that the fundamental rights and freedoms of individuals concerned are protected and a fair balance is struck when they benefit from these privately-owned genetic genealogy databases in their investigations. At this point, it must be remembered that the ECtHR underlined that the use of modern scientific techniques in the criminal justice system and investigating cold cases should not impede the right to privacy of the individuals concerned; instead, a carefully-considered balancing test should be conducted in these cases.[86] Therefore, at the dawn of the rise of IGG in Europe, we must urgently open the debate to a wider public to understand if and how this technique could be deemed lawful and legitimate in the EU legal framework, so we can ensure that it is not implemented at the expense of our fundamental rights and freedoms.

## 5. Conclusion

IGG has emerged recently as a ground-breaking investigation tool allowing investigators to solve cases that may otherwise remain unsolved. However, like any other advancement in criminal investigations, this tool too did not come without controversy. Several ethical and legal concerns have been voiced against this novel investigation technique, specifically related to the privacy interests of the users of genetic genealogy databases and their

85 Folketinget, *supra* n. 15.

86 S. and Marper, *supra* n. 17, §112; Gaughran, *supra* n. 18, §93. *See, also* Aycaguer, *supra* n. 18, §34; Trajkovski and Chipovski, *supra* n. 17, §51.

biological relatives. With the European LEAs demonstrating a growing interest in IGG, it has become crucial to address these concerns to ensure that IGG is not used at the expense of our fundamental rights and freedoms. Although the ECtHR and CJEU jurisprudence on DNA retention provide an overview of the requirements needed to strike a fair balance between society's interest in the fight against crime and the fundamental rights and freedoms of the individuals concerned, IGG may cause these legal restraints to be practically circumvented, since LEAs benefit from access to databases beyond the national forensic DNA databases. Accordingly, further research appears to be necessary to understand whether and how the legality and legitimacy of this novel investigation technique can be ensured in Europe. Therefore, this chapter concludes by calling for urgent action on this matter.

CHAPTER

# V

# Legal Assessment of Digital Political Campaigning: The Right to Free Elections

**Keyomars Khaleghi**[1]

https://doi.org/10.26116/a87p-hn80

1 PhD Researcher, Tilburg Institute of Law, Technology, and the Society (TILT), Tilburg Law School (TLS), Tilburg University

## 1. Introduction

The formal investigation of the Information Commissioner's Office in the UK regarding the use of data analytics for the microtargeting of political adverts during the EU referendum (Brexit) became the largest investigation of its type. The investigation involved online social media platforms, data brokers, analytics firms, academic institutions, political parties and campaign groups.[2] On the academic side, in their recent study on the effects of social media on election outcomes in the United State, Fujiwara, Müller & Schwarz conclude that 'Twitter's relatively liberal content may have persuaded voters with moderate views to vote against Donald Trump'.[3] These are examples of the rising evidence that using algorithmic technologies in election campaigns[4] affects free elections.

Generally, political research has acknowledged the profound impact of emerging digital technologies and social media on politics. Some scholars have assessed this impact as the reordering of politics by digital technologies[5] and the structural impact of social media on the political environment[6]. In specific, critical concerns have been raised in political communication research regarding the manipulation of emerging digital technologies and social media in political campaigning. Algorithmic technologies have enabled political campaigners to segment audiences and send targeted messages on social media platforms quickly and at very low cost.[7] Scientific evidence proves that the capabilities of social media platforms and digital technologies have been utilised in political campaigns during elections in order to influence their results.[8] From a functional perspective, 'technique' is the most appropriate concept for describing the use of

2 Denham, E. (2018). Investigation into the use of data analytics in political campaigns. *A report to Parliament. Information Commissioner's Office*, *6*, p.7.

3 Fujiwara, T., Müller, K., & Schwarz, C. (2024). The effect of social media on elections: Evidence from the united states. *Journal of the European Economic Association*, *22*(3), 1495-1539, p. 1495.

4 '"[E]lection campaign" refers to a set of systematic and organised efforts and actions aimed to influence the voters' decision making', Recommendation (CoE) CM/Rec (2022)12 of the Committee of Ministers to member States of 6 April 2022 on electoral communication and media coverage of election campaigns.

5 *See* Amoore, L. (2023). Machine learning political orders. *Review of International Studies*, *49*(1), 20-36.

6 *See* on concerns about the extent to which people consume political content via social media, Reuning, K., Whitesell, A., & Hannah, A. L. (2022). Facebook algorithm changes may have amplified local republican parties. *Research & Politics*, *9*(2), 20531680221103809.

7 *See* Stier, S., Bleier, A., Lietz, H., & Strohmaier, M. (2020). Election campaigning on social media: Politicians, audiences, and the mediation of political communication on Facebook and Twitter. In *Studying politics across media* (pp. 50-74), p. 50-51. Routledge. Also *See* on the centrality of social media in political campaigning in Europe, Daniel, W. T., & Obholzer, L. (2020). Reaching out to the voter? Campaigning on Twitter during the 2019 European elections. *Research & politics*, *7*(2), 2053168020917256.

8 *See* Bradshaw, S., & Howard, P. N. (2018). Challenging truth and trust: A global inventory of organ-

digital technologies on social media for political campaigning purposes. The most powerful techniques that political campaigners have used extensively since 2015 on social media platforms are profiling,[9] disinformation,[10] echo chambers,[11] social bots[12] and microtargeting[13] (hereinafter 'the digital campaigning techniques').

There is a consensus among political communication scholars that the practice of political campaigning has entered a new era. The four characteristics of political campaigning in the new era are: (1) organisational and strategic dependency on digital technology and big data, (2) personalised and microtargeted campaign messages with persuasive approach, (3) internationalisation of the campaigns in terms of campaign actors and communication with voters, and (4) networked communication approach to voters, especially through social media platforms.[14] In this environment, the campaign actors are both human and automated.[15] Big data, highly-developed algorithmic technologies and the pervasive use of social media have become the key dynamic elements of political campaigning. In addition to being led by campaign professionals, many

ized social media manipulation. *The computational propaganda project*, *1*, 1-26. <http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/07/ct2018.pdf>.

9 'Profiling means 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', Article 4(4) of 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) OJ L 119/33.

10 '[M]isleading, false or scurrilous information (content) served to people with the intention of influencing political discourse and elections, a phenomenon come to be labelled "fake news" or "online disinformation".' Opinion 3/2018 of The European Data Protection Supervisor of 19 March 2018 on Online Manipulation and Personal Data (EDPS), <https://edps.europa.eu/sites/edp/files/publication/18-03-19_online_manipulation_en.pdf>.

11 Echo chambers are the operation of algorithms and automated recommender systems that in which users only see pieces of information that confirm their own opinions or match their profile. The Committee of Experts on Internet intermediaries (MSI-NET) Council of Europe DGI (2017)12 of March 2018 on Algorithms and Human Rights: Study on the Human Rights Dimensions of Automated Data Processing Techniques and Possible Regulatory Implications, 30, <https://edoc.coe.int/en/internet/7589-algorithms-and-human-rights-study-on-the-human-rights-dimensions-of-automated-data-processing-techniques-and-possible-regulatory-implications.html>.

12 'Social bots are algorithmically controlled accounts that emulate the activity of human users but operate at much higher pace (e.g. automatically producing content or engaging in social interactions), while successfully keeping their artificial identity undisclosed', ibid.,32.

13 'Online political microtargeting is a type of personalised communication that involves collecting information about people and using that information to show them targeted political advertisements. Borgesius, F. J. Z., Möller, J., Kruikemeier, S., Fathaigh, R. Ó., Irion, K., Dobber, T.,... & De Vreese, C. (2018). Online political microtargeting: Promises and threats for democracy. *Utrecht Law Review*, *14*(1), 82-96, p. 82.

14 Roemmele, A., & Gibson, R. (2020). Scientific and subversive: The two faces of the fourth era of political campaigning. *New Media & Society*, *22*(4), 595-610, p. 597.

15 *See* n.3, *Supra*.

digital political campaigns are citizen-initiated, and they rely more on non-professional and peer-to-peer masses of citizens.[16] On social media, identifying political campaigns and advertisements is becoming more complex. In fact, distributing the political words of a campaigner over social media is largely labelled as expressing a personal opinion even though they might be organised as a political advertisement or campaign. In practice, emerging digital technologies have enabled individual campaigners to purchase bots and pay people to spread their message over social media.[17] In addition, AI has adopted a significant role in the improvement of digital political campaigns.[18] New generations of AI persuasion technologies, specifically persuasive dialogue systems, can be used to create AI applications that can engage users in an automated dialogue and persuade them to believe something.[19] In short, the trends in digital political campaigning are:[20]

- superior targeting and audience segmentation;[21]
- cross-device targeting;[22]
- the increasing use of psychographic analysis;[23]
- growing use of AI for campaigning techniques;[24]
- use of AI to generate content automatically;
- effective delivery of campaigns through social media.

16 Römmele, A., & von Schneidmesser, D. (2016). Election campaigning enters a fourth phase: the mediatized campaign. *Zeitschrift für Politikwissenschaft*, *4*(26), 425-442, p. 427.

17 *Digital campaigning: Increasing transparency for voters* (The Electoral Commission), Last updated: 9 June 2021, p. 7.

18 *See*, Watts, C. *Artificial Intelligence is Transforming social media. Can American Democracy Survive?*, (The Washington Post), September 2018, <https://www.washingtonpost.com/news/democracy-post/wp/2018/09/05/artificial-intelligence-is-transforming-social-media-can-american-democracy-survive/?noredirect=on&utm_term=.7eb9b9a23203>.

19 APS maintains a model for the user which enables the system to choose good moves in the dialogue to persuade. *See* Hunter, A. (2018). Towards a framework for computational persuasion with applications in behaviour change. *Argument & Computation*, *9*(1), 15-40.

20 Bartlett, J., Smith, J., & Acton, R. (2018). *The future of political campaigning*. Demos.

21 *See*, Tambini, D. (2018) *Social Media Power and Election Legitimacy*. In: Tambini, D. & Moore, M. (eds.) Digital dominance: the power of Google, Amazon, Facebook, and Apple. Oxford University Press, New York, NY, pp. 265-293

22 *See*, Chester, J., & Montgomery, K. C. (2017). The role of digital marketing in political campaigns. *Internet Policy Review*, *6*(4), 1-20.

23 *See*, Chen, L., Gong, T., Kosinski, M., Stillwell, D., & Davidson, R. L. (2017). Building a profile of subjective well-being for social media users. *PloS one*, *12*(11), e0187278.

24 *See*, Brundage, M., Avin, S., Clark, J., Toner, H., Eckersley, P., Garfinkel, B.,... & Amodei, D. (2018). The malicious use of artificial intelligence: Forecasting, prevention, and mitigation. *arXiv preprint arXiv:1802.07228*.

The 2016 presidential election in the US and the investigations surrounding the mishandling of millions of Facebook users' data for the purpose of campaigning throughout the presidential election[25] led to the demand for research on the legal dimensions of using algorithmic technologies in election campaigning. In Europe, legal discussions about the digital campaigning techniques have mostly been formed around political microtargeting and protecting personal data under data protection law. A wider approach in legal research on the digital campaigning techniques has been initiated within the framework of human rights law. This chapter is intended to suggest a human rights law approach to research on the digital campaigning techniques. This is an approach based on the right to free elections.[26] Section 2 reviews the literature on legal aspects of using the digital campaigning techniques. Section 3 presents the reasons that the right to free elections approach is necessary for examining the impact of the digital campaigning techniques. The subsections in section 3 explore the key requirements for the right to free elections approach: the systematic and holistic view and precision in the assessment of impact. Section 4 looks at capacity improvement and the challenges presented by the right to free elections approach.

## 2. Literature Review

The literature review initially provides an outline of the research on identifying key legal risks related to using algorithmic technologies in political campaigning. Even though the outline includes some references to the US election environment, it focuses on political campaigning in Europe. Then, an overview of the legal research on the risks related to the use of specific algorithmic techniques is provided, focusing on microtargeting and disinformation in political campaigning.

### 2.1 Identification of Key Legal Risks

Bennett provides a general picture of the data-driven election campaigning environment and identifies the key related legal risks. He identifies legal risks like voter surveillance, privacy and data protection. He compares the general legal context of the data-driven electoral campaigning in North America with Europe, and he observes that the data protection regulations in Europe create effective protection against the impacts of data-

25 Kang, C., Rosenberg, M. & Frenkel, S. *Facebook Faces Broadened Federal Investigations Over Data and Privacy*, (The New York Times), 2 July 2018, <www.nytimes.com/2018/07/02/technology/facebook-federal-investigations.html>.

26 Article 3 of Protocol No. 1 to the European Convention on Human Rights, 'The High Contracting Parties undertake to hold free elections at reasonable intervals by secret ballot, under conditions which will ensure the free expression of the opinion of the people in the choice of the legislature.'

driven election campaigning.[27] Dobber et al. explore the barriers to and facilitators for targeting specific groups in elections with tailored messages through digital tools used by Dutch political parties. They echo Bennett's finding that there are important differences between the US and Europe in terms of the regulatory frameworks that influence how messages tailored by digital tools are used. They believe 'the relatively strict Dutch data protection law' regulates the processing of political preferences as 'sensitive personal data' for political behavioural targeting.[28] Manheim & Kaplan examine data-driven election campaigning with a focus on the capabilities artificial intelligence (AI) created for political campaigners, mainly in the US, and the risks that using AI would pose to privacy and democracy. They compare the regulatory frameworks of the US and the EU with regard to controlling the risks of abusing AI in election campaigning as well. They believe that the use of data analytics does not distort election processes per se. As such, they take the position that there is a difference between the legitimate and illegitimate use of data and algorithms in election campaigning.[29]

Brkan points out the impact of using AI on the freedom of elections in the EU. She has a different view regarding the sufficiency of the existing regulatory framework concerning the use of algorithmic technologies in electoral campaigning. She identifies critical risks to European democracy due to the potential of AI technology for convincing and manipulating voters in elections. She concludes that the EU regulatory frameworks should strike a balance between the right to free elections and freedom of expression, media freedom and media pluralism.[30] Brkan argues that digital political campaigning affects voters' right to privacy, personal data, freedom of information (by limiting voters' access to information about all political parties) and the right to free elections. She clarifies that manipulating data-driven campaigns presents a risk to public interests, specifically freedom of elections and democracy. As such, data protection regulations that cover private interests of data subjects cannot sufficiently protect freedom of elections. She argues that assessing the adverse effects of data-driven political campaigning on the right to free elections from a legal perspective requires discussing free elections in terms of more political values. Her assessment is that unjustifiable interference with freedom of expression or information provision to voters by manipulative online political campaigns does not seem to lead to the violation of the fundamental right to free elections automat-

27 Bennett, C. J. (2016). Voter databases, micro-targeting, and data protection law: can political parties campaign in Europe as they do in North America?. *International Data Privacy Law*, *6*(4), 261-275.

28 Dobber, T., Trilling, D., Helberger, N., & De Vreese, C. H. (2017). Two crates of beer and 40 pizzas: the adoption of innovative political behavioural targeting techniques. *Internet Policy Review*, *6*(4), 1-25, p. 7.

29 Manheim, K., & Kaplan, L. (2019). Artificial intelligence: Risks to privacy and democracy. *Yale JL & Tech.*, *21*, 106-188, p. 138.

30 Brkan, M. (2019). Artificial intelligence and democracy: The impact of disinformation, social bots and political targeting. *Delphi*, *2*, pp. 66-68.

ically, however.[31] Similarly, Helberger et al. claim that political advertising needs an entirely different regulatory framework. They suggest drawing lessons from experiences with regulating unfair commercial advertising for designing EU regulations on digital political advertising and online political targeting. They recognise that the critical combination of detailed knowledge about voters and their behaviours has generated a new level of persuasion, creating the risk of this leading to unfair forms of manipulation that undermine voters' autonomy.[32] In addition, they claim that social media platforms are important drivers for 'increasing commercialisation of political advertising and are blurring the lines between commercial and political advertising'.[33]

### 2.2 Microtargeting as a Campaigning Technique

Borgesius et al. take a multidisciplinary approach to discussing the promises and threats of online political microtargeting and explore how policymakers in Europe can intervene in controlling the risks of online political microtargeting from both legal and social science perspectives. They believe that the risks of manipulating online political microtargeting in the EU can mostly be controlled through data protection law. They conclude that if the risks of online political microtargeting were to materialise, they would threaten democracy. In their view, the 'risks should not be overstated'.[34] Based on their assessment, online political microtargeting may have less influence in Europe than in the US because of differences in the legal and electoral systems. They extend this optimistic view by arguing that the influence of online political advertising on voters has limits.[35] Ronan Ó Fathaigh et al. examine the risk of manipulating political microtargeting and the dissemination of disinformation by foreign actors during elections in Europe. They argue that the advances in AI technology enable foreign actors to run propaganda more effectively, more efficiently and less overtly than traditional propaganda.[36] They assess that combining microtargeting techniques with novel propaganda techniques could affect elections in Europe. To conclude, they reiterate that the enforce-

31 Brkan, M. (2020). EU fundamental rights and democracy implications of data-driven political campaigns. *Maastricht Journal of European and Comparative Law*, *27*(6), 774-790, pp. 780-781 & 786.

32 Helberger, N., Dobber, T., & de Vreese, C. (2021). Towards unfair political practices law: Learning lessons from the regulation of unfair commercial practices for online political advertising. *J. Intell. Prop. Info. Tech. & Elec. Com. L.*, *12*, 273-296, p. 273.

33 Ibid., 278.

34 *See* n.12, 82 *Supra.*

35 Ibid., 96.

36 Ó Fathaigh, R., Dobber, T., Zuiderveen Borgesius, F., & Shires, J. (2021). Microtargeted propaganda by foreign actors: An interdisciplinary exploration. *Maastricht Journal of European and Comparative Law*, *28*(6), 856-877, pp 863-865.

ment of data protection law in Europe could help mitigate the threat of microtargeted foreign propaganda,[37] although they suggest some potential additional rules.[38]

Dobber et al. discuss the matter of insufficiency of the data protection law for regulating political microtargeting in Europe. They explain that, currently, there is no specific regulatory framework for political microtargeting in Europe and only general rules apply to this field, including data protection law, privacy law, freedom of expression and sector-specific rules for political advertising. They argue that the General Data Protection Regulation (GDPR) framework is necessary for controlling the risk of political microtargeting, but that it is not sufficient.[39] In order to clarify the insufficiency of the GDPR, they claim that EU lawmakers did not consider the specific context of microtargeting when they were drafting the GDPR. In addition, the right to freedom of expression and democracy values, as the different dimensions of microtargeting in election campaigning, are not protected under the GDPR.[40]

### 2.3 Disinformation as a Campaigning Technique

Hoboken et al. provide an overview of the altered media landscape as a result of the utilisation of digital disinformation. They believe this new media landscape, which is controlled by a relatively small group of internationally-operating internet service providers, has a major impact on democratic processes. Further, they claim that the process of disseminating disinformation has changed, and digital techniques, such as microtargeting, social bots, and trolls, create additional complexity in the effects of disinformation on democratic processes.[41] In fact, microtargeted political disinformation can engage targeted voters more effectively, and the use of automated bots makes identifying the source of disinformation extremely difficult.[42] This indicates the interrelationships between the digital campaigning techniques, such as social bots, microtargeting and disinformation, and how this interrelation may lead to the digital campaigning techniques having increased effects.

37 Ibid., 867.

38 Ibid., 869-871.

39 On how GDPR could assist to regulate digital political campaigns, Dobber et al. say 'it is harder for political parties to buy data about people. And in most countries in Europe, it is impossible to access voter registration records.' Dobber, T., Ó Fathaigh, R., & Zuiderveen Borgesius, F. J. (2019). The regulation of online political micro-targeting in Europe. *Internet Policy Review*, *8*(4), 1-20, p. 7.

40 Ibid., 7.

41 van Hoboken, J., Appelman, N., Ó Fathaigh, R., Leerssen, P., McGonagle, T., van Eijk, N., & Helberger, N. (2019). The legal framework on the dissemination of disinformation through Internet services and the regulation of political advertising, p. 11.

42 Ibid., 21.

Bayer et al. argue that major disinformation campaigns over the past four years illustrate interference in elections and referenda. They provide recommendations on legislative and policy measures to protect democracy, the rule of law and fundamental rights against the impact of disinformation. Their examination on the impact of disinformation focuses on privacy, human dignity, autonomy and freedom of expression,[43] whereas Hoboken & Ó Fathaigh focus more on setting standards for freedom of expression for dealing with disinformation. The latter clarify that the European Democracy Action Plan[44] requires dealing with disinformation through actions that create an empowering environment for freedom of expression, such as:

> funding projects to support deliberative democratic infrastructures, strengthening media freedom and media pluralism, and strengthening empowerment of citizens to make informed decisions through strengthening media literacy.[45]

They believe this framework is aligned with the approach of the ECtHR that lays out that states not only have a duty of non-interference with freedom of expression, but that they also have an obligation to put appropriate legislative and administrative frameworks in place to create an enabling environment for effective pluralism.[46]

It can be concluded that multiple legal scholars have examined the legal risks posed by the use of algorithmic technologies in election campaigning as in general and within specific contexts. They identified several risks, including data protection, privacy, human dignity and freedom of expression. However, there is still little research on the use of algorithmic technologies in election campaigning from a right to free elections perspective. Brkan highlights the public interest aspect of electoral campaigning, and she suggests examining the use of algorithmic technologies in electoral campaigning from a right to free elections perspective. She also emphasises that assessing the impact of data-driven political campaigning on the right to free elections requires considering the respective political elements. Brkan's ideas question the sufficiency of data protection law for controlling the risk of using algorithmic technologies in political campaigning, and she raises the need to integrate the right to free elections in legal research.

---

43 Bayer, Judit and Holznagel, Bernd and Lubianiec, Katarzyna and Pintea, Adela and Schmitt, Josephine and Szakács, Judit and Uszkiewicz, Erik, Disinformation and Propaganda: Impact on the Functioning of the Rule of Law and Democratic Processes in the EU and Its Member States – 2021 Update (April 24, 2021). Available at SSRN: https://ssrn.com/abstract=4092020

44 Communication (EU) COM/2020/790 of European Commission to the European Parliament of 3 December 2020 on the European democracy action plan.

45 van Hoboken, J., & Fathaigh, R. Ó. (2021). Regulating Disinformation in Europe: Implications for Speech and Privacy. *UC Irvine J. Int'l Transnat'l & Comp. L.*, *6*, 9, p. 35.

46 Ibid.

In order to integrate the right to free elections into the examination of the digital campaigning techniques, two key questions need to be answered:

i. Why is the right to free elections approach necessary?
ii. What are the requirements of assessing the digital campaigning techniques from the right to free elections perspective?

The following sections provide an overview of why the right to free elections should be considered as an essential approach to evaluating digital campaigning techniques and the key components of this approach. Next, the essential requirements of the right to free elections approach are examined: the systematic and holistic view, and precision.

## 3. The Right to Free Elections Approach

Some of the risks that arise from the political campaigning techniques to democracy have been highlighted in EU policies. In the European Democracy Action Plan, the EU demands more transparency in political advertising and communication, specifically transparency in the enforcement of relevant rules, audits, access to non-personal data, the restriction of microtargeting and psychological profiling in political communication.[47] Based on this view, ensuring political democracy is directly related to political communication and it is an essential component of free elections. Protecting personal data was the starting point of legal research into the impact of the digital campaigning techniques, yet the direct relation between democracy and political communication indicates a need to move beyond the protection of personal data. In general, the right to free elections is the most relevant legal framework for assessing digital campaigning techniques in terms of their impact on free elections and political democracy.[48] In fact, no other fundamental right directly concerns the impact and consequences of using digital campaigning techniques for free elections and political democracy. The right to free elections includes states' positive obligations to ensure free expression for voters regarding their choice of legislature over elections, considering the realities of election campaigns.[49] In addition, ensuring political democracy as an issue of

47 n.43, 2.1 *Supra*.

48 ECtHR in Guide on Article 3 of Protocol No. 1 expresses, 'In order for the rights guaranteed by Article 3 of Protocol No. 1 to be effective, their protection cannot remain confined to the candidature itself. The election campaign thus also falls within the scope of the provision.' *Guide on Article 3 of Protocol No. 1*, § 91 (Council of Europe/European Court of Human Rights, 2022).

49 ECtHR analysed in The Communist Party of Russia and Others v. Russia case, 'The next question is thus whether the State was under any positive obligation under Article 3 of Protocol No. 1 to ensure that media coverage by the State-controlled mass-media was balanced and compatible with the spirit of "free elections", even where no direct proof of deliberate manipulation was found.' ECtHR concludes, '…where the case concerns the extent of the State's positive obligations,

public interest[50] cannot be secured solely under the personal data protection legal framework. The personal data protection legal framework is essentially intended for protecting the processing of personal data of an identified or identifiable natural person and not the public dimension of manipulating data that concerns the protection of free elections and democracy.[51] Furthermore, a considerable amount of data used in digital campaigning techniques is not personal data, so it is beyond the scope of the data protection legal framework. In addition, the legal assessment of the digital campaigning techniques requires considering how digital techniques are used to communicate with voters and not only the methods of collecting and processing personal and non-personal data. This in turn demands assessment of the impact of each of the digital campaigning techniques on the right to free elections.

In examining the impact of the digital campaigning techniques on the right to free elections, consideration of the interdependence and interrelation of this fundamental right with other fundamental rights is critical. In particular, consideration of the interdependence between the right to freedom of expression (freedom of political expression through digital political campaigning techniques) and the right to free elections, and striking a balance between these fundamental rights is essential.[52] The impact of using digital technologies for psychological targeting in election campaigns on intellectual privacy[53] of voters as an important part of the right to privacy also has a direct relation with the right to free elections. In fact, voters' distorted autonomy in relation to elections due to psychological targeting may affect their right to free elections.

The legal examination of the effects of digital campaigning on the right to free elections needs to embrace its political nature. Accordingly, the outcomes of research in the

and that the State is only required to take those measures which are "reasonably available"', The Communist Party of Russia and Others v. Russia, no. 29400/05, §123, 19 June 2012.

50 Birch argues that the essence of a democratic electoral model is 'the notion that the citizenry elects people to public office in order to serve its ends, to serve the public interest'. Birch, S. (2011). *Electoral malpractice*. Oxford University Press, USA, p. 27.

51 On the scope of data protection, *See* Amann v. Switzerland [GC], no. 27798/95, ECHR 2000-II, § 65. Even though processing of political opinions of a person is prohibited under Article 9 of GDPR, but this limitation only relates political opinions and does not cover the consequences of manipulating political data that may affect free elections.

52 'Free elections and freedom of expression, particularly freedom of political debate, together form the bedrock of any democratic system. The two rights are inter-related and operate to reinforce each other: for example, as the Court has observed in the past, freedom of expression is one of the "conditions" necessary to "ensure the free expression of the opinion of the people in the choice of the legislature".' Bowman v. The United Kingdom, no. 24839/94, § 42, 19 February 1998.

53 Intellectual privacy, 'typified by a person's interest in privacy of thought and mind, and the development of opinions and beliefs, Koops, B. J., Newell, B. C., Timan, T., Skorvanek, I., Chokrevski, T., & Galic, M. (2016). A typology of privacy. *U. Pa. J. Int'l L.*, *38*, 483, p. 567. Also See, Richards, N. (2015). *Intellectual privacy: Rethinking civil liberties in the digital age*. Oxford University Press, USA.

political communication sphere are used as evidence to assess the impact of the digital campaigning techniques on free and fair elections. Practically, the legal elements of the right to free elections are used as the criteria to assess whether the evidence from political communication on the impact of the digital campaigning techniques[54] has an adverse impact on the right to free elections. Also, political communication theories that respect digital campaigning and social media are essential for understanding the functionality[55] of each of the digital campaigning techniques. However, two key challenges are associated with using political communication research outcomes as evidence for legal assessment. First, the validity of evidence from political communication is not static due to constant enhancement of algorithmic technologies and data ecosystems. This leads to constant changes in the validity of evidence. As a result, legal research cannot rely on current evidence, and it needs to examine new political communication evidence based on the complexity and growing capacity of algorithmic technologies and data-driven ecosystems. Second, the results of political communication research on the impact of the digital campaigning techniques and the role of social media in election campaigning have sometimes been contradictory.[56] This has led to confusion in selecting the right evidence for legal assessment. This issue is not a phenomenon in political science.[57] Social science theorists have explored it and recommended solutions based on validity and reliability concepts.[58]

54 For instance, the research by Orestis Papakyriakopoulos et al. show that data mining techniques enable information collection about a person's general opinion, party preferences and other non-political characteristics on social media. In addition, application of algorithmic technologies, enable profiling to identify whom and how to micro target and influence voters in elections. They discuss the ethical and political implications of profiling and microtargeting on social media for German political system. This could be among the evidence for legal assessment of the impact of profiling and microtargeting on the right to free elections. Papakyriakopoulos, O., Hegelich, S., Shahrezaye, M., & Serrano, J. C. M. (2018). Social media and microtargeting: Political data processing and the consequences for Germany. *Big Data & Society*, *5*(2), 205395171881184.

55 'Functionality' in this text means, the tasks that the digital campaigning techniques are capable to perform.

56 For instance, Baldwin-Philippi argues that 'data-driven practices have been much more productive at mobilising action, like getting out the vote and improving donation rates, than at persuasive goals of getting someone to support a candidate.' In contrast, Bradshaw & Howard believe social media is being used to manipulate and deceive the voting public and to undermine democracies. Baldwin-Philippi, J. (2019). Data campaigning: Between empirics and assumptions. *Internet Policy Review*, *8*(4), 1-18, p.2. Bradshaw, S., & Howard, P. N. (2018). Challenging truth and trust: A global inventory of organized social media manipulation. *The computational propaganda project*, *1*, 1-26, p. 21. For further reading on the contradictory research outcomes, *See* Bennett, C. J., & Lyon, D. (2019). Data-driven elections: implications and challenges for democratic societies. *Internet policy review*, *8*(4).

57 See Beer, F. A. (1993). Validities: A political science perspective. *Social Epistemology*, *7*(1), 85-105.

58 See Drost, E. A. (2011). Validity and reliability in social science research. *Education Research and perspectives*, *38*(1), 105-123.

### 3.1 A Holistic and Systematic View

The functionality of the digital campaigning techniques is entirely dependent on the political campaigning environment as explained earlier and on the elements of the political campaign, namely big data, highly-developed algorithmic technologies and pervasive use of social media. Consequently, legal research on the digital campaigning techniques that does not consider the environment and these elements would be imprecise. Algorithmic technologies and social media act as systematically interconnected elements in the digital campaigning techniques. In this systematic function, social media has become the infrastructure that algorithmic technologies function in to operationalise the digital campaigning techniques. As such, the success of the digital campaigning techniques depends heavily on the efficient use of algorithmic technologies and social media.

The need for a systematic approach in research on algorithmic technologies has been identified by some legal and non-legal scholars. Bodó et al. believe that there are:

> pressing challenges about how to stay in control of digital environments which are increasingly co-habited and controlled by opaque algorithmic agents, weaving non-transparent personalized-experience-cocoons around individual users.

They conclude that this non-transparency creates what they call an 'algorithmic control crisis'.[59] They explain that among the factors are a 'control crisis,' which they define as the lack of 'systematic research into the normative implications of algorithmic control'.[60] Woolley & Howard explain the necessity of taking a systematic view, specifically in a political communication context. They believe that, together, 'social media, political bots, and the Internet of things enable computational propaganda.'[61] They define computational propaganda as 'the assemblage of social media platforms, autonomous agents, and big data tasked with the manipulation of public opinion'.[62] They clarify that, within the computational propaganda system, social media platforms operate as an infrastructure that 'autonomous agents, equipped with big data about our behaviour collected from the Internet of Things' use.[63] A systematic approach

59 Bodo, B., Helberger, N., Irion, K., Zuiderveen Borgesius, F., Moller, J., van de Velde, B.,... & de Vreese, C. (2017). Tackling the algorithmic control crisis-the technical, legal, and ethical challenges of research into algorithmic agents. *Yale JL & Tech.*, *19*, 133, p. 139.

60 Ibid., 139.

61 Woolley, S. C., & Howard, P. N. (2016). Automation, algorithms, and politics| political communication, computational propaganda, and autonomous agents—Introduction. *International Journal of Communication*, *10*, 9, p. 4886.

62 Ibid., 4886.

63 Ibid.

to the digital campaigning techniques entails mapping and understanding the related algorithmic technologies and data-driven ecosystems, also considering the capabilities of social media as the infrastructure of political campaigning. This systematic view is a requirement for taking the right to free elections approach to assessing the impact of the digital campaigning techniques because the right to free elections demands transparency in terms of the impact and consequences of the systematic function of social media and algorithmic technologies. Since the right to free elections is directly connected to democracy,[64] the systematic approach to the impact assessment of the digital campaigning techniques necessarily includes the consequences for democracy as well. This highlights the difference between the data protection law approach and the right to free elections approach, which concentrates on the consequences of the use of data and technology on free elections.

Taking a holistic view of the digital campaigning techniques requires considering the whole cycle of political campaigning and the actors involved in it. In their recommended model on the democratic use of technology in elections and political campaigning, Neudert & Howard suggest addressing four key stakeholder groups, namely civil society, government, the digital industry[65] and political parties. They also suggest targeting the whole campaign cycle, including preparation, the campaigning period and the post-campaign evaluation.[66]

### 3.2 Need for Precision

Assessing the impact of the digital campaigning techniques on the right to free elections requires precision. This precision is necessary for describing the use and functionality of the digital campaigning techniques, and to perform a risk assessment of them. The need for precision derives from the sensitivity of identifying the right balance between the application of the right to freedom of expression within political campaigning and the limits that might be applied on the campaigns to protect the right to free elections. An imprecise impact assessment would interfere with the right to the freedom of expression for political campaigners and political actors. It would also adversely affect policies encouraging a favourable environment for innovation and digital transformation in the

64 ECtHR's interpretation of Article 3 of Protocol No.1 is, 'According to the Preamble to the Convention, fundamental human rights and freedoms are best maintained by "an effective political democracy". Since it enshrines a characteristic principle of democracy, Article 3 of Protocol No. 1 (P1-3) is accordingly of prime importance in the Convention system.' Mathieu-Mohin and Clerfayt v. Belgium, no. 9267/81, § 47.

65 Digital industry actors encompass political advertising service providers and digital technology service providers.

66 Neudert, L. M., & Howard, P. (2019). Ready to vote: elections, technology and political campaigning in the United Kingdom <https://oxtec.oii.ox.ac.uk/publication/ready-to-vote/>, p. 4.

EU, specifically the Digital Decade Policy Programme 2030[67] and the Digital Strategy.[68] To guarantee the necessary precision, a risk-based approach for assessing the impact of the digital campaigning techniques on the right to free elections would be sufficient.

With regard to the impact assessment method, a combination of Guidance on Human Rights Impact Assessment of Digital Activities and Human Rights, Democracy,[69] and Rule of Law Impact Assessment (HUDERIA)[70] would be efficient. These methods would improve the quality of impact assessments of the digital campaigning techniques. These impact assessment methods are used to assess the severity of the impacts of the digital campaigning techniques on the right to free elections, considering scope, scale and irremediability of the impacts. Primary and secondary legal sources that relate to the right to free elections provide a legal framework to define the criteria for identifying the scope and assessing the impact of the digital campaigning techniques.[71] These impact assessment methods were originally developed for assessing the impact of business operations on human rights, but they are sufficiently flexible to be used for human rights impact assessments with a wider approach as well. The methods use parameters for assessing impact severity. Impact assessment methods should embrace: (1) the state's duty to protect the right to free elections; (2) corporate responsibility to respect the right to free elections; and (3) the right holder's access to remedy for violating the right to free elections.

67 European Declaration on Digital Rights and Principles for the Digital Decade 2023/C 23/01, OJ C 23, 23.1.2023.

68 European Commission digital strategy, Next generation digital Commission, 30.6.2022 C(2022) 4388 final [2022].

69 *Human rights impact assessment of digital activities*, The Danish Institute for Human Rights (November 2020), <https://www.humanrights.dk/publications/human-rights-impact-assessment-digital-activities>.

70 *Human Rights, Democracy, and the Rule of Law Assurance Framework for AI Systems: A proposal prepared for the Council of Europe's Ad hoc Committee on Artificial Intelligence*, The Alen Turing Institute (March 2021), <https://rm.coe.int/huderaf-coe-final-1-2752-6741-5300-v-1/1680a3f688>.

71 Among the instruments that are used for identifying the scope and assessing the impact of the digital campaigning techniques are: the Act concerning the election of the representatives of the Assembly by direct universal suffrage, EU Code of Practice on Disinformation, Council of Europe Recommendation CM/Rec (2020)1 of the Committee of Ministers to Member States on the human rights impacts of algorithmic systems.

## 4. Capacity Improvement and Challenges

The proposed approach to assessment of the impact of the digital campaigning techniques would increase capacity in the following areas:

1. Multidimensional assessment of the impact of the digital campaigning techniques – The holistic and systematic characteristics of the suggested approach and the integration of the whole cycle and actors in digital political campaigns would allow for a multidimensional and reliable impact assessment of the digital campaigning techniques from a legal perspective.
2. Providing criteria for identifying harmful digital political campaigns – The proposed approach would define the right to free elections as a criterion for identifying and assessing harmful and illegal digital political campaigning. This comes from the capacity of the right to free elections to protect free elections against abusing data and algorithmic technologies. This indicates that the criteria for identifying and assessing commercial campaigning are different from political campaigning and advertising. Obviously, this approach would not be a replacement for the protection of personal data within the respective legal frameworks.
3. Preciseness in the assessment of risks – The proposed approach would add precision to the risk assessment of the digital campaigning techniques. This would be beneficial for defining and striking a balance between the right to free elections and the right to freedom of expression.
4. Evidence-based impact assessment – The proposed approach demands using the results of political communication research as evidence for the legal impact assessment of the digital campaigning techniques. This would increase the precision and reliability of the impact assessment.
5. Evaluation of the effectiveness of the existing regulatory framework – By assessing the scope and severity of the impact, the proposed approach could effectively assist in evaluating the effectiveness of the existing regulatory framework for protecting free elections against the use of the digital campaigning techniques.
6. Clarity on regulatory strategy – The proposed holistic and systematic approach would assist in examining how the current EU regulatory framework could be improved for protecting free elections and for evaluating the existing regulatory framework if a specific regulation is needed regarding the digital campaigning techniques.
7. Reconsidering irremediability – As was explained earlier, the right to free elections is directly connected to effective democracy. Therefore, the systematic approach to the impact assessment of the digital campaigning techniques requires identifying the consequences for democracy as well. Identifying the consequences for democracy requires reconsidering legal remedies to include the adverse impact on democracy as

well. For instance, violation of the right to free elections by using the digital campaigning techniques impacts a free and fair election. This impact may require annulment of an election or preventing a long-term influence in power by a political party as a remedy.[72]

8. More clarity on dimensions of legal research – The proposed approach suggests integration of public interest in legal research on the digital campaigning techniques. This would add a new dimension to the existing legal research, which focuses on private interest in terms of data protection and privacy.
9. Flexibility in legal research – The proposed risk-based approach provides flexibility in legal research so it can shift from a solely binary system of legal or illegal acts to assessing potential risks of illegality regarding the use of data and algorithmic technologies in the digital campaigning techniques.
10. Transparency within the policy and the EU's regulatory framework – The policies of the EU are intended to promote democracy within the context of the Democracy Action Plan. On the other hand, they are also intended to encourage an internal market and to incentivise a business-driven digital transformation in line with the Digital Decade Policy Programme 2030 and the Digital Strategy. Thus far, these two directional strategies have been defined within the Digital Services Act, the Proposal for Artificial Intelligence Act, the Proposal for Data Governance Act, the proposed Regulation on the Transparency and Targeting of Political Advertising, and the Proposal for a Regulation on Privacy and Electronic Communications as a regulatory framework. The proposed approach would enable the assessment of both sides of the EU policies, and they have been equally promoted in the area of political communication.

There are also challenges associated with the right to free elections approach.

1. The right to free elections under Article 3 of Protocol No. 1 of the European Convention on Human Rights only concerns the choice of legislature. The European Court of Human Rights has clarified that referenda and presidential elections do not fall under the scope of Article 3 of Protocol No. 1 in its case law.[73] This may affect the

72 The importance of the effectiveness of remedies, where a violation of Article 13 of ECHR in conjunction with the right to free election takes place, was explained in the Petkov and Others v. Bulgaria case. The ECtHR extends the limits of effective remedy for serious breaches of the right to freedom of election that affect the outcomes to annulment of an election result as a remedy. However, the level of the impact of utilising the digital campaigning techniques could move beyond the outcome of a specific election result. Political campaigners could use the capabilities of the algorithmic technologies for long term influence in power. This situation would require reconsidering the remedy for violation of the right to free elections.

73 *See* Cumhuriyet Halk Partisi v. Turkey (dec.), no. 48818/17, §§ 33 & 38; and Moohan and Gillon v. the United Kingdom (dec.), nos. 22962/15 and 23345/15, § 40.

scope of research on the digital campaigning techniques from the right to free elections perspective. However, the nature of referenda and presidential elections as democratic processes is not different from parliamentary elections. As such, the aforementioned limitations do not affect the suggested research approach. The scope of legal research from the right to free elections perspective would also encompass the impacts on voters and the fundamental right of voters to freely express their opinions, disregarding the form or purpose of the democratic process as a parliamentary election, presidential election or referendum.

2. As explained in section 3, there are two challenges associated with using the outcomes of political communication research as evidence for legal assessment. First, the validity of political communication evidence is constantly changing owing to continuous improvements in algorithmic technologies and data ecosystems. Second, the results of political communication research on the impact of the digital campaigning techniques and the role of social media in election campaigning have sometimes proven contradictory. Social science theorists have explored this issue and recommended solutions based on validity and reliability concepts.

## 5. Conclusion

Scholars have examined the legal dimensions of using algorithmic technologies in election campaigning, namely microtargeting and disinformation. They have identified some risks posed by data protection, from privacy, human dignity and the freedom of expression perspectives. This chapter suggested examining the impact of the digital campaigning techniques from the right to free elections perspective. The necessity of examining the impact of the digital campaigning techniques on the right to free elections derives from the insufficiency of other relevant legal disciplines. A systematic and holistic view and precision in impact assessment are key requirements for a right to free elections approach. The systematic view requires seeing algorithmic technologies and social media as systematically interconnected elements in the digital campaigning techniques. In this systematic function, social media acts as the infrastructure that algorithmic technologies work on to operationalise the digital campaigning techniques. The holistic view of the digital campaigning techniques requires the inclusion of the entire cycle of political campaigning and the actors in this cycle. Precision is necessary for describing the functionality of the digital campaigning techniques, and risk assessment of them. The need for precision derives from the importance of identifying the right balance between the application of the right to freedom of expression within political campaigning and the limits that might be applied in campaigns for protection of the right to free elections. A legal examination of the impact of digital campaigning on the right to free elections would be associated with the political nature of this right.

As such, the outcomes of research in political communication and digital social media spheres could be used as evidence to assess the impact of the digital campaigning techniques on free and fair elections from a legal perspective.

CHAPTER

# VI

# A Rights-Based Defence of Cognitive Manipulation by Artificial Intelligence

**Aimen Taimur**[1]

https://doi.org/10.26116/7463-ct34

1 PhD Researcher, Tilburg Institute of Law, Technology, and the Society (TILT), Tilburg Law School (TLS), Tilburg University.

# 1. Introduction

## 1.1 Background

The Orwellian concept of 'Thoughtcrime',[2] illustrated as an exaggerated form of state control and destruction of personal liberty, has become a fast-approaching reality in light of the twenty-first century, artificial-intelligence-powered new technologies. The problem that needs to be addressed is the violation of one of the three absolute rights:[3] freedom of thought, by way of the obstruction of natural cognitive reasoning with AI-designed thought manipulation. Tracking online user activity to create a landscape of the *forum internum*[4] and the use of this profiling to infer personal vulnerabilities in order to facilitate cognitive manipulation has a profoundly negative impact on how we fundamentally understand normative agency, democracy and liberty.[5] This calls for an analysis of the robustness and relevancy of human rights protections intended to save natural persons against the breach of the absolute right to think freely, which encapsulates the defence against having unexpressed thoughts accessed, read, manipulated, criminalised[6] or suppressed.[7] The understanding of freedom of thought and for it to be breached through cognitive hacking, the abstractness of the *forum internum* and the resultant intersectional dimension of the manipulation that may occur pose an unprecedented assault on our fundamental rights. It needs to be determined whether politico-legal reactions to thought conditioning are in line with the supranational human rights obligations of the Universal Declaration of Human Rights, specifically those outlined in Article 18 (freedom of thought and conscience) and Article 19 (freedom of opinion and freedom to hold opinions without interference) of the International Covenant on Civil and Political Rights. Moreover, it is still yet to be seen if and subject to what limitations international human rights law and domestic regulations protect individuals from pervasive AI thought manipulation techniques.

2 George Orwell's Novel 1984 explains 'Thoughtcrime' as the criminalisation of dissenting opinions inside one's mind, contrary to the interests of the establishment.

3 Human Rights Guide, 'Absolute Rights' (https://www.cilvektiesibugids.lv/en/themes/human-rights-restrictions/absolute-rights#:~:text=For%20example%2C%20the%20internal%20aspect,restricted%20and%20is%20considered%20absolute)

4 *Forum Internum* directly translates to 'internal forum', but it is legally used to refer to the private inner space of the mind.

5 Mill, J. S. 'On Liberty', Chapter 2: 'Of the Liberty of thought and discussion', (1859)

6 Farahany, N. A. (2012). Incriminating thoughts. *Stan. L. Rev.*, *64*, 351.

7 McCarthy-Jones, S. (2019). The autonomous mind: The right to freedom of thought in the twenty-first century. *Frontiers in Artificial Intelligence*, *2*, 19.

## 1.2 Thought Control in Real Time

It has been recorded that the primary caretakers of Fundamental Rights, the governments of democratic states, have been party to the exploitation of the manipulative powers of AI to sway political opinion.[8] The risk to the privacy of one's mental space became excruciatingly evident after the emergence of the neuropolitical campaigning documented during the 2016 Cambridge Analytica scandal.[9] In this case, data that was collected through social media surveillance by an analytics firm was used for behavioural micro-targeting through Facebook advertisements to condition the political opinion of civilians before the Brexit vote and during the US Presidential elections. Cambridge Analytica was engaged as a consulting firm for the campaigns of two well-known Republican candidates in the US elections of 2016 before receiving legal notice.[10] The company used data science techniques to identify the exact social group to target to ensure their political messaging was effective. The target demographic was determined to be primarily swing voters who still offered the possibility of being persuaded to support the Republican party or Trump supporters.[11] The company largely used Facebook to gather information for precise and focused political advertising. There are rumours that the information utilised to create psychographic profiles was obtained via a different app called 'This Is Your Digital Life App,'[12] which connected users and their friends' Facebook profiles after the app was downloaded and served as a conduit for data extraction from Facebook databases. Facebook issued a statement after the leak became public, reiterating its disclaimer and updating its security protocols.[13]

A few years later, similar concerns were raised about the political misuse of social media platforms for campaigning by political parties in the 2019 Canadian elections.[14] Links have been made between mood alteration and deliberate online content curation

8 Bradshaw, S., & Howard, P. N. (2018). Challenging truth and trust: A global inventory of organized social media manipulation. *The computational propaganda project*, *1*, 1-26.

9 Boldyreva, E. (2018). 'Cambridge Analytica: Ethics and Online Manipulation with Decision Making Process', 91 – 102 (December 2018) <https://www.researchgate.net/publication/330032180_Cambridge_Analytica_Ethics_And_Online_Manipulation_With_Decision-Making_Process>.

10 Smith, A. (2018). There's an open secret about cambridge analytica in the political world: It doesn't have the 'secret sauce'it claims. *Business Insider, March*, *21*, 2018.

11 Lewis, P., & Hilder, P. (2018). Leaked: Cambridge Analytica's blueprint for Trump victory. *The Guardian*, *23*.

12 The Guardian, (2018) 'Revealed: Aleksandr Kogan collected Facebook users' direct messages', <https://www.theguardian.com/uk-news/2018/apr/13/revealed-aleksandr-kogan-collected-facebook-users-direct-messages>.

13 Grewal, P. (2018). Suspending Cambridge Analytica and SCL Group from Facebook. *Facebook Newsroom*, *16*.

14 Bradshaw, J. (2018). *Securing Canadian Elections: Disinformation, Computational Propaganda, Targeted Advertising and What to Expect in 2019*. Canadian International Council (CIC).

in Facebook's 2012 experiment to check the effect of the material displayed on individual News Feeds and its influence over the users' emotional states.[15] There has also been concern over the commercial use of the access and reading of the *forum internum* to sell products, as can be seen in Facebook's alleged admission of having the ability to read emotions and micro-target advertisements to insecure and mentally vulnerable teenagers in real-time.[16]

It goes without saying that because this is such a widespread and systematic practice, it has practically become the norm. It is not adequately protected by either the EU e-Commerce directive, which only stipulates that notice must be given in the event of a targeted advertisement and does not forbid attention harvesting.[17] The same is true of the Unfair Commercial Practices directive,[18] which only expresses concern about aggressive commercial practices that may compel customers to make decisions but does not provide additional guidance on whether attention harvesting and targeted advertisements are sufficiently 'aggressive' to warrant legal action. With its explicit declaration of complete protection for vulnerable minors,[19] the draft Digital Services Act[20] aims to prevent the escalation of commercial exploitation through targeted advertising, but it also fails to provide a comprehensive framework for the protection of attentional privacy. This raises a more significant question, as highlighted by the AI Act, regarding the broader implications of targeted advertisements. While the Digital Services Act recognises that these advertisements have the potential to manipulate behaviour, particularly when they capture attention and exploit the vulnerabilities of minors, it does not go far enough in providing comprehensive legal protections. Specifically, it fails to address the wider effects of this type of exploitation on larger demographic groups, leaving a significant gap in the regulatory framework. Because of this, the design of the digital

15 Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*(24), 8788-8790.

16 Levin, S. (2017). Facebook told advertisers it can identify teens feeling 'insecure'and 'worthless'. *The Guardian*, *1*.

17 Directive 2000/31 – Legal aspects of information society services, in particular electronic commerce, in the Internal Market ('Directive on electronic commerce')

18 Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council ("Unfair Commercial Practices Directive"), Section 2 – Art 8&9

19 Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (Text with EEA relevance), Art 28.2 & Art 24

20 European Commission, 'Questions and Answers: Digital Services Act', 25 April 2023, <https://ec.europa.eu/commission/presscorner/detail/en/QANDA_20_2348>

marketplace[21] is now obviously asymmetrical in terms of power dynamics, making it exploitative to consumers.

## 2. Jurisprudential Grounding

The human mind is the last frontier of privacy and this has been discussed philosophically in the times preceding artificial intelligence in the shape of the concept of the *forum internum*. The proposition of the *forum internum* was first brought to light in the context of religious thought.[22] Human rights theorists have also defined *forum internum* broadly as an internal private sphere or the 'inner realm of the mind'.[23] Entitlement to this private space was considered an essential part of the manifestation of religious thought, which consisted of deciding how one was spiritually aligned, with no interference or control allowed by external entities, such as the state.[24]

Charles Taylor's social theory provides a template to explain how the relationship between the outdated conceptualisation of the *forum internum* and how a necessary reclamation of the concept in the context of manipulative AI may be deconstructed.[25] As a global first-order practice, most systems are based on the presumption that individual actions are a result of autonomous, individual reasoning. The validity of this reasoning comes from the authority one has over reaching a conclusion through independent cognition without any interference from external forces. Modern democracies rely heavily on this concept, wherein the act of casting a vote is what gives politicians the mandate to hold public office. The vote is symbolic of free will and highlights the ability of each individual to decide who may be entrusted with the resources and who they would want to be ruled by.[26] Similarly, the vote also acts as a social contract which communicates positive intent which is necessary to establish contractual validity. Intent has significant legal importance within criminal law too, since it can change the intensity of sentencing. For example, a homicide may be classified as murder, which has far worse consequences than manslaughter, which is known to have relatively lenient sentencing

21 Bignull, H. (2023), 'Deceptive Patterns Exposing the Tricks Tech Companies use to Control You'.

22 Taylor, P. M. (2005). The scope of the forum internum beyond religious choice. In *Freedom of Religion: UN and European Human Rights Law and Practice* (pp. 115–202). chapter, Cambridge: Cambridge University Press.

23 Roberts, C. K. (2016). Interpreting 'freedom from'religion: A step too far?. <https://legalresearch.blogs.bris.ac.uk/2016/06/interpreting-freedom-from-religion-a-step-too-far/> accessed 15th March 2023.

24 Stenlund, M., & Slotte, P. (2018). Forum internum revisited: Considering the absolute core of freedom of belief and opinion in terms of negative liberty, authenticity, and capability. *Human Rights Review*, *19*(4), 425-446.

25 Taylor, C. (1985). Social Theory as Practice. Cambridge University Press 91 – 115.

26 Peter, F. (2009). *Democratic legitimacy*. Routledge.

guidelines.[27] This is what the element of *mens rea* is, which requires intent to commit a crime to correspond with the *actus reus* which is the actual commission of the crime. For liability to be found and sentencing to be carried out, both elements need to be proven beyond reasonable doubt. Hence, psychological autonomy is the common underpinning within organised frameworks of representation and justice.

The second-order practice is what brings the discrepancy between the practical implementations of laws and the attack on cognitive freedom to the fore, by highlighting the legal interpretative limitation on the freedom of thought and the inability of other available legislative protections to preserve cognitive freedom. Our collective social needs and understanding of mental privacy are not congruent with the way that the laws provide protection. This suggests a critique of the human rights framework, with a focus on the freedom of thought, which is inadequately applied because the freedom to think is ultimately unprotected. The correct formulation of appropriate legal reform is necessary to eliminate deficiencies in the present structure. The identification and acknowledgement of the importance of the inner realm of the mind predates the technological revolution, which the legislative frameworks must reflect in the defence they provide.

## 3. Protective Rights

The evolving sophistication of AI and its widespread adoption make the demographic for psychological manipulation broader than ever before. As the AI influence on independent natural choice increases, the human rights protections which are intended to defend free cognitive functioning must adapt in parallel to build a fortress around individuals and their sense of self.[28] This has been highlighted recently as a matter of scholarly concern within the general discussion on the psychological impact of technology on the human experience at large.[29] It has already been shown that AI has the ability to intrude upon and shape individual thoughts.[30] The Council of Ministers of the European Union[31] have also cautioned against the potential manipulation of natural cognition by AI and have pushed for further research on the matter.

27 McAuley, F. (1982) 'Mens Rea: A Legal Philosophical View' *Irish Jurist* 84-104.

28 Koops, B. J., Newell, B. C., Timan, T., Skorvanek, I., Chokrevski, T., & Galic, M. (2016). A typology of privacy. *U. Pa. J. Int'l L.*, *38*, 483.

29 Hattenstone, S. (2023). Tech guru Jaron Lanier:'The danger isn't that AI destroys us. It's that it drives us insane.'. *The Guardian*, *23*.

30 Wilson, H., Rauwolf, P., & Bryson, J. J. (2021). Evolutionary psychology and artificial intelligence: The impact of artificial intelligence on human behaviour. *The SAGE handbook of evolutionary psychology: Applications of evolutionary psychology*, 333-351.

31 Council of Europe, 'Declaration by the Committee of Ministers on the manipulative capabilities

So far, the rights-based inquiry into the interference with the *forum internum* has been grounded in the understanding of two rights specifically: the right to privacy and the right to freedom of thought. However, the traditional understanding of these rights does not coherently encapsulate the right to mental self-determination.[32] Arguably, the human rights remedies set in place lack the nuance and sufficient adaptability to accommodate flexible interpretations to sufficiently protect normative agency[33] from external alteration by third parties via AI. Therefore, it remains to be seen whether and how the existing human rights regime can eliminate the risk to mental privacy and the impending danger of unexpressed thoughts being controlled.[34]

The second option is the creation of a new set of rights called 'neurorights'. Neurorights are laws that guard against cognitive and emotional manipulation of an individual's central nervous system by artificial intelligence in order to effectively protect the rights to a private life, personality and sense of agency, among others. The main function that neurorights must play is to go beyond fundamental human rights.[35] The critique that this approach usually attracts is the importance of clearly identifiable laws which can be adopted and implemented with ease. To avoid overcrowding of legislation, it has been recommended to make use of existing laws to facilitate the protection offered under neurorights. Yet one of the largest issues that remains is the ascertaining of the ability and traceability of fault in case of breaches.

## 4. Ascertaining Liability

The use of AI-powered technologies to influence human behaviour leads to the question of how the law, at least in theory, would deal with the extent of manipulation to potentially shape intolerant[36] views and to ascertain the responsibility of harm caused by individuals who have had interference with their reasoning faculties through this new recalibration of power over the mind. It is known that social media algorithms are

of algorithmic processes' (Adopted by the Committee of Ministers on 13 February 2019 at the 1337th meeting of the Ministers' Deputies).

32 Alegre, S. (2017). Rethinking freedom of thought for the 21st century. *European Human Rights Law Review*, *3*(13), 221-33.

33 Brownsword, R. (2017). Law, liberty and technology. *The Oxford handbook of law, regulation and technology*, 41-68.

34 Gartner, M. (2022). Regulatory Acknowledgment of Individual Autonomy in European Digital Legislation: From Meta-Principle to Explicit Protection in the Data Act. *Eur. Data Prot. L. Rev.*, *8*, 462.

35 European Science Media Hub, 'Neurorights: Do our brains need to be protected by legislation?', <https://sciencemediahub.eu/2023/11/08/neurorights-do-our-brains-need-to-be-protected-by-legislation/>.

36 Müller, K., & Schwarz, C. (2021). Fanning the flames of hate: Social media and hate crime. *Journal of the European Economic Association*, *19*(4), 2131-2167.

designed to produce higher levels of consumer engagement,[37] which results in personalised content being pushed according to individual interests, which produces online thought bubbles that act as echo chambers limited to information that garners individual user attention, depriving the user of alternative knowledge to form informed opinions. Foreseeably, this could create an environment that impedes the free flow of information, subconsciously framing the thoughts of unsuspecting users.

Recently, the members of the European Parliament (MEPs) and the Council reached a provisional agreement on the Artificial Intelligence Act, which is a historic development that represents a comprehensive regulatory framework for guaranteeing the safety of AI in Europe. The main goals are to protect democracy, fundamental rights and environmental sustainability while simultaneously encouraging innovation to establish Europe as a leader in artificial intelligence. The agreement includes strong bans on particular AI applications that are thought to be dangerous for citizens' rights and democratic principles. Applications that are prohibited include sensitive characteristic-based biometric categorisation systems, untargeted face image scraping for recognition databases, emotion recognition in educational and work environments, behaviour-based social scoring and artificial intelligence systems that manipulate behaviour to subvert free will or take advantage of weaknesses.[38]

In the EU, following extensive policy discussion, a four-tier risk assessment model has been used to broadly categorise the threats posed by artificial intelligence in the current draft of the EU AI Act (European Union's Artificial Intelligence Act).[39]

Artificial intelligence poses a risk, but it is categorised as an 'unacceptable risk'[40] in the four categories of unacceptable risk, high risk, limited risk and minimal risk. The Act mandates that certain practices must be expressly 'prohibited' as a stringent industry standard in response to the identification of unacceptable risk. However, in actuality, the proscribed ban cannot be carried out if distinct lines are drawn around cognitive behavioural manipulation, which leads to a lack of specificity when assigning blame to a particular party after a crime has been committed. There is not a single, clearly-provable charge with prerequisites for prosecution. The primary legal issue this raises is whether the protections for the average citizen are still applicable in cases where they are the target of deliberate cognitive manipulation that results in behavioural modifi-

37 Zuboff, S. (2023). The age of surveillance capitalism. In *Social theory re-wired* (pp. 203-213). Routledge.

38 EU Parliament Press Release,' Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI' <https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai>

39 Artificial Intelligence Act. (21 April 2021). "Proposal for a regulation of the European Parliament and the Council laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act) and amending certain Union legislative acts." *EUR-Lex – 52021PC0206* https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELLAR:e0649735-a372-11eb-9585-01aa75ed71a1.

40 Ibid, Art 5

cation. Furthermore, the legal responsibility for tech companies and data analytics firms is also still obscure, giving them an ever-widening window of opportunity to avoid accountability as algorithmic manipulation techniques get more sophisticated and less simple to explain over time.[41]

Article 5 of the EU AI Act was formulated with the intention of governing AI manipulation to forestall potential deleterious consequences. Nevertheless, the effective execution of this legislative framework proves to be challenging, attributable to the presence of ambiguous terminologies and the unclear delineation of manipulative techniques. As a result, Article 5 has been critiqued on multiple fronts, namely [1] regarding its perceived deficiency in lucidity [2], insufficiency in protective provisions [3] and apparent inability to realise its specified objectives [4]. These impediments are symptomatic of the dynamic evolution of AI technologies, novel legal complexities with scarce precedent within extant legal literature and research, but also of the interdisciplinary nature innate to AI governance.[42]

There is also the matter of the Act's potential lack of effectiveness against other low-ranking, significant risks if it cannot establish an appropriate accountability framework for the highest form of presentable risk, which is its self-identified risk. The straightforward identification is invalidated because the Act lacks an outline that codifies the perception of the risk, its manifestations and the expected punishment in the event that the risk materialises. Even though the AI Act is still being reviewed before it is ultimately adopted, there are as of yet issues regarding its interpretation that must be resolved in order for it to be effective in controlling and supervising artificial intelligence innovation and reducing ethical risks that could have detrimental effects on society as a whole. Additionally, for the new AI Act to enhance the current EU digital legal framework, symbiosis between the pre-existing fundamental rights obligations and their application in the context of new technologies is required.

Adequate redressal and the ability to identify the practical contours of the need that the legislation is meant to fill are essential preconditions for drafting sufficiently protective legislation. Since it will provide an analysis of the harm that the law seeks to remedy, this purposeful approach to the protection of cognitive freedom is the best course of action. Two primary arguments aim to address the gap that exists in all current laws,[43] as suggested by the literature in the field as a type of legal barrier against the

41 Franklin, M., Ashton, H., Gorman, R., & Armstrong, S. (2022, May). Missing mechanisms of manipulation in the EU AI Act. In *The International FLAIRS Conference Proceedings* (Vol. 35).

42 Zhong, H. (2023). Regulating AI manipulation: Applying Insights from behavioral economics and psychology to enhance the practicality of the EU AI Act. *arXiv preprint arXiv:2308.02041*.

43 Sevastianova, V. N. (2023). Trademarks in the age of automated commerce: Consumer choice and autonomy. *IIC-International Review of Intellectual Property and Competition Law*, *54*(10), 1561-1589.

effects of manipulative AI:[44] the establishment of new rights (neurorights)[45] or the discovery of safeguards within the current human rights legislation to expressly preserve cognitive freedom.

That being said, it is necessary to precisely identify the scope of the problem before evaluating the merits of any proposed solutions. The damage needs to be chronologically reconstructed to its origin in order to comprehend the thought manipulation process. The English common law practice of determining fault through the establishment of causation can be utilised to accomplish this task in the manipulation case. Drawing a chain with a manipulated thought as the final link – which might or might not lead to an action – is the first step in this process. The chain begins to form once the restrictions of the *forum internum* have been bypassed and the mind's private inner space has been entered. Harvesting attention is the simplest way to explain this access. 'Reading' the unexpressed thoughts and opinions that an individual may hold is the next step after this accessing is completed. Behaviour analysis and psychographic profiling have already demonstrated that this is a reality with a fair degree of accuracy.

Psychographic profiling has historically been used to examine individual behaviours within the framework of gathering marketing data and establishing a connection with commonplace activities to estimate the supply and demand of goods in a precise manner for commercial objectives. Either statistical data from tracking purchasing patterns or a generalised market analysis from a competitive landscape[46] were used as the input for this data. However, the volume of data stored by digital spaces has increased significantly due to the scale of data collection following the popularisation of social media and the digitisation of databases.

The second phase involves manipulation, which can take many different forms. Some of these have been discussed in previous concepts, such as recommender systems that "nudge" people toward particular decisions. Apart from recommender systems, another frequently-employed deceptive technique is the use of dark patterns, which are intended to fool people into making decisions online that they otherwise would not have made consciously.

Furthermore, these patterns are thought to only work on people when they are unaware of the influence being applied to them, and coercion is not considered a concealed influence, while manipulation is. The study findings on dark pattern identification indicate significant differences between the proposed designs. Few users recognised the pre-selection nudge, coerced permission and dark patterns based on deception

44 Alegre, S. (2023). We don't need'neurorights' we need to know the existing law. *The Financial Times*, 17-17.

45 The Neurorights foundation posits this thesis in their mission statement, <https://neurorights-foundation.org/>.

46 Solomon, M. R. (2004). Consumer psychology. *Encyclopedia of applied psychology*, *1*, 483-492.

techniques, but the majority recognised the high-demand/limited-time message and 'confirmshaming' for, for example, trick questions, loss-gain framing and hidden information. Although these results pertain only to a specific application of the dark pattern and are not generalisable to the category,[47] they might suggest that certain dark patterns are intrinsically more difficult to detect. The question of whether consent can be obtained from a user without granting them the freedom to develop a well-founded, explicit permission arises in light of the evolution of manipulative practices in digital spaces and the absence of explicit regulation surrounding them. This indicates a suppression of people's ability to make decisions on their own because people are already being manipulated and forced into making decisions in digital spaces, often without even realising it.

The next step after manipulation is the potential for thought suppression, which can put an end to the capacity to hold opposing views. The complete subservience of the mind to whatever the manipulating party wants it to believe is ensured by suppressing the innate human capacity for original thought. Another problem that arises from this complete reduction of free thought is isolating people from objective reality and putting them in a position where they are incapable of making independent decisions based on reason and data. The last step is the criminalisation of particular ideas and viewpoints, which creates a society in which expressing dissent or engaging in alternative thought is frowned upon.

It is still not set in stone whether the stipulated human rights protections simply provide a non-binding compulsion on a state to regulate AI operations that breach the *forum internum* or whether they would pave a clear way for citizens to demand an effective remedy from the court for breach of their right to mental privacy. Further, how would such litigation technically progress, considering the new age interpretation of the freedom of thought which would cover unexpressed thoughts? This requires the evaluation of the domestic legal compliance in providing individuals with the right to mental privacy and the potency of oversight mechanisms in ensuring adequate safeguards for the protection of the *forum internum*.[48] Lastly, the relevance of the findings must be assessed in the context of imminent technological singularity with new AI-powered tools being developed and put on open markets for mass usage. This assessment should involve studying the points of alignment and divergence between the interna-

47 Bongard-Blanchy, K., Rossi, A., Rivas, S., Doublet, S., Koenig, V., & Lenzini, G. (2021, June). " I am Definitely Manipulated, Even When I am Aware of it. It's Ridiculous!"-Dark Patterns from the End-User Perspective. In *Proceedings of the 2021 ACM Designing Interactive Systems Conference* (pp. 763-776).

48 Petkoff, P. (2012). Forum internum and forum externum in canon law and public international law with a particular reference to the jurisprudence of the European Court of Human Rights. *Religion & Human Rights*, 7(3), 183-214.

tional standards of thought protection currently in place and the predicted impacts on individual rights. These impacts arise from inadequate administrative attitudes that enable convenient AI-based manipulation while fostering the strategic use of such technologies.[49]

## 5. Future-Proof Legislation

With the fear of pervasive manipulative AI outpacing the law, there are also academic narratives of legislating afresh to control the arbitrary violation of cognitive sovereignty.[50] Although this has been done mostly considering the nonconsensual neuro-technical interventions of the human mind, even if a similar solution is employed to regulate manipulative AI, prospective legislation has drawbacks which revolve around getting outdated with the rapid developments of the tech industry.[51] Since the research issue is relatively new, the lack of case law highlights a dearth of guidance for policymakers to use as a foundation for effective regulation. Moreover, the radically advancing nature of AI would outdate the regulations put down to control them since AI algorithms keep changing post-production. Deep Learning and autonomous AI decision-making also raises a question about the relevancy of the strict liability model.[52] It has not yet been answered if it would be possible to control AI with law reform and whether said reform would equally hold states, corporations and private individuals responsible while giving weightage to the complex decision-making ability of AI.

Alternatively, arguments have been put forth from the neo-liberal anti-regulation perspectives, which dictate that restricting technological development is contrary to human growth.[53] With the ninth UN Sustainable Development Goal being the fostering of innovation, there is overwhelming support for technological advancement, without foresight for the necessary regulatory regimes being put in place to deal with potential fundamental rights violations. This has also been the reason behind the lack of support for work to prevent AI's violation of the human mental ecosystem. Various new developments within artificial intelligence such as the highly sophisticated ChatGPT and its upgraded version GPT-4 have already hit the market and are publicly-accessible AI

49 Krishnamurthy, V. (2021). AI and Human Rights Law. *Artificial Intelligence and the Law in Canada (Toronto: LexisNexis Canada, 2021)*.

50 Bublitz, C. (2015). Cognitive Liberty or the International 83 Human Right to Freedom of Thought.

51 Christopher Reed, 'How to make bad law: lessons from cyberspace' Mod. Law Rev. 73, 903– 932 (2010)

52 Bygrave, Lee A., Machine Learning, Cognitive Sovereignty and Data Protection Rights with Respect to Automated Decisions, University of Oslo Faculty of Law Research Paper No. 2020-35 (29th October 2020)

53 Vargas, L. G. (2017). Do They Want to Regulate Online Profiling?. *Canadian Journal of Law and Technology*, *15*(1).

platforms. Through such sophisticated user-friendly programs, not just corporations but also individuals can develop working codes for apps, websites, data-mining programs and online advertisements, which means that the global audience, which may be exposed to techniques for behavioural manipulation powered by artificial intelligence, is limitless.[54] Even though these developments are examples of large language models, it is expected that such programs will allow an unregulated mass of individuals to create workable algorithms for their own ulterior gains. The introduction of such a resource is difficult to regulate since it gives rise to the question of whether the creators of such open-access AI should be given liability in case of unethical usage of their services or the individuals who maliciously rely on these services. The current technique for avoiding liability employed by most platforms is either content regulation or responsibility waivers incorporated within the terms and conditions. The efficacy of both these techniques is not absolute, and they might not even work for the creative liberty that open artificial-intelligence programs may give users in the coming years.[55]

## 6. Conclusion

The idea of liberty has gained a lot of attention over time, but cognitive liberty is largely invisible and has been sidelined in discussions regarding ethics of artificial intelligence due to the obscurity of what it may entail and the difficult traceability of its violation. Most laws target the protection of personal authority in the public sphere while ignoring the restoration of rights in the private sphere. Existing literature indicates an abundance of macro discourse against technological solutionism specifically in the areas of criminal profiling, e-courts and judicial due process, yet the micro-discourse of autonomous cognitive reasoning that occurs in the human mind is severely overlooked. This is evident from the absence of a detailed regulative model even from international forums which have identified the threat of AI's thought manipulation.[56] The theoretical discourse on regulating these issues is hindered by the presence of laws that, while existing, are often rendered ineffective due to vague definitions and a lack of clarity. The concept of threats to freedom of thought and mental privacy is not entirely new, as George Orwell's *1984* famously introduced the notion of Thoughtcrime as a chilling

54 Andrew Meyers, 'AI's Power of Political Persuasion', Stanford University Human Centered Artificial Intelligence (27th February 2023) <https://hai.stanford.edu/news/ais-powers-political-persuasion>.

55 Helmore, E. (2023). We are a little bit scared': OpenAI CEO warns of risks of artificial intelligence. *The Guardian*, *17*.

56 ISO/ IEC DTR 24368, Information Technology -Artificial Intelligence – Overview of ethical and societal concerns <https://www.iso.org/standard/78507.html?browse=tc> accessed 17th March 2023.

example. However, the legal frameworks currently in place to address these concerns lack the necessary clarity and effectiveness. This inadequacy is further compounded by the unprecedented precision of modern, non-invasive methods of intruding into the human mind, highlighting the urgent need for more effective and clearly-defined legal protections. Although protection against violation of mental integrity has been codified in Article 3 of the European Charter of Human Rights, it is purely in terms of clinical rather than non-medical external mass conditioning. A viable solution to restrict this specific incursion and prevent a global human rights crisis is still pending. It is clear, however, that an update to our private liberty with the inclusion of cognitive liberty needs to be made in order to conserve human intelligence in an era dominated by artificial intelligence.

PART

# III

# RECONFIGURING THE LEGAL PARADIGM FOR THE PUBLIC SECTOR

CHAPTER

# VII

# Fit for Purpose? The Role of Consent in EU Data Protection Law in Light of Very Large Online Platforms' Processing of Personal Data

**Ana-Maria Hriscu &**

**Eleni Kosta**[1]

https://doi.org/10.26116/wpjm-jf97

1 PhD researcher and Professor at the Tilburg Institute for Law, Technology and Society (TILT) at Tilburg University, the Netherlands.

## 1. Introduction

Consent has been seen as a means for data subject empowerment in European data protection since the adoption of the Data Protection Directive (DPD).[2] Consent in the DPD is seen as a tool for the empowerment of the 'participatory right of informational self-determination'.[3] During the review of the European data protection framework, the consent of the data subject was at the centre of the debates. One of the key objectives of the European Commission was the examination of 'ways of clarifying and strengthening the rules on consent'.[4] The General Data Protection Regulation (GDPR)[5] retained the DPD's definition of consent almost intact,[6] and it was complemented with specific rules that related to consent in Article 7 GDPR and to the consent of children in relation to information society services in Article 8 GDPR. Consent is only one of the six lawful grounds for the processing of personal data. However, it is broadly used as a legal basis for the processing of personal data, especially by private companies. Concurrently, arguments are increasingly being made on the actual validity of the provided consent in these contexts. The aim of this chapter is to reflect on the role of consent in European data protection law and to explore the question of whether consent can fulfil its empowering role given very large platforms' data processing practices.

Section 2 of this chapter recounts the role that consent has been given in EU data protection law as a means of empowering individuals to exercise control over their personal data, and the crucial role transparency obligations play in this respect. It then

2 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. O.J. (L 281/31) 24.11.1995.

3 Mayer-Schönberger, V. (1997). *Generational development of data protection in Europe* in *Technology and privacy: The new landscape* (Agre, P. & Rotenberg, M. (Eds.) The MIT Press 1997) 235. Rossangel et al. have characterised the consent of the individual for the processing of their personal data as the 'genuine expression of the right to informational self-determination': Roßnagel, A., Pfitzmann, A., & Garstka, H. J. (2001). *Modernisierung des Datenschutzrechts: Gutachten im Auftrag des Bundesministeriums des Innern* (pp. 1-9). Berlin: Bundesministerium des Innern.

4 European Commission, *Communication from the Commission to the European Parliament, the Council, the Economic and Social Committee and the Committee of the Regions on a comprehensive approach on personal data protection in the European Union* (2010) COM (2010) 609 final, 04.11.2010, 9.

5 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), O.J. (L 119/1) 27.4.2016.

6 Article 2(h) of the Data Protection Directive defined consent as 'any freely given specific and informed indication of his wishes by which the data subject signifies his agreement to personal data relating to him being processed'. The GDPR defines consent in Article 4(11) as 'any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her'.

outlines the requirements for valid consent and transparent data processing set out in EU data protection law. Section 3 introduces three instances of very large online platforms'[7] processing of personal data on the basis of consent. The three selected processing activities are common among very large platforms that offer a variety of products and services: 1) the processing of personal data, including the setting of cookies for targeted advertising; 2) the processing of special categories of personal data; and 3) the processing of children's personal data. These processing activities, taken from privacy policies, are analysed for their compatibility with requirements regarding consent and transparency. The processing activities analysed are not representative of how all online platforms process personal data based on consent. Rather, they are illustrative of the issues surrounding meaningful consent and transparency. Section 4 maps out the main criticisms of consent in the online context as found in the literature. These, together with the findings from section 3, form the basis for our reflection in the final section of this chapter. This reflection concerns whether consent can fulfil its purpose of empowering individuals in the online context, and what other legal interventions may be necessary to safeguard individuals' right to data protection.

## 2. Consent and Transparency – Tools for Empowerment in EU Data Protection Law

### 2.1 The Rationales of Consent and Transparency

According to Article 5(1)(a) of the GDPR, personal data shall be processed 'lawfully, fairly and in a transparent manner in relation to the data subject'. The principle of lawfulness requires that data controllers rely on at least one of the six legal bases enumerated in Article 6 GDPR when processing personal data. The separate but 'intrinsically linked and interdependent'[8] principles of fairness and transparency require that processing is fair and that data subjects are provided with information about the processing. For a data controller to rely on the consent legal basis, consent must be valid. This means it must be informed, specific, freely given and that it must represent an unambiguous expression of the data subject's wishes.[9] Since its first encapsulation in data protection

7 Namely Google, Meta, and Microsoft. This research uses the definition of very large platforms provided by Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) O.J. (L 277/1) 19.10.2022. Article 23 defines very large platforms as those 'which provide their services to a number of average monthly active recipients of the service in the Union equal to or higher than 45 million'.

8 European Data Protection Board (EDPB), *Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR)*, para 219.

9 Article 4(11) GDPR.

regulation, consent has been framed as a means through which individuals can be empowered to exercise control over the collection and uses of their personal data. Already in the German Federal Data Protection Act of 1977, one of the first pieces of data protection legislation in Europe, consent was introduced as grounds for data processing, in addition to the other legal provisions based on which the processing of personal data was allowed.[10] Soon after in the 1983 Population Census case in the German Constitutional Court, the right to informational self-determination was recognised as 'the capacity of the individual to determine in principle the disclosure and use of his/her personal data on the basis of the concepts of dignity and self-determination in the German Constitution'.[11] While this notion of informational self-determination, with consent as one of its core tenets, did not result in a broad adoption within the constitutional orders of Member States,[12] it was influential in the evolution of EU data protection law nonetheless.[13]

Interestingly enough, 'informed consent' was first introduced in the early proposals for a data protection directive as part of the rights of the data subject.[14] However, even at the time, the industry feared it would become too cumbersome to obtain informed consent from data subjects.[15] After a lengthy legislative process, the consent of the data subjects was introduced as a ground for lawful processing, defined in Article 2(h) of the Data Protection Directive.[16] This structure remained in the GDPR, enhanced with rules on conditions for consent[17] and consent of children.[18] After the Lisbon Treaty, the EU Charter of Fundamental Rights (hereinafter 'CFR' or 'Charter'),[19] recognised the right to

10 Kosta, E. (2013). *Consent in European data protection law* (Vol. 3). Martinus Nijhoff Publishers, 50.

11 De Hert, P., & Gutwirth, S. (2009). Data protection in the case law of Strasbourg and Luxemburg: Constitutionalisation in action. In *Reinventing data protection?* (pp. 3-44). Dordrecht: Springer Netherlands.

12 For instance, in the Netherlands. *See e.g.* Brouwer, E. (2008). *Digital borders and real rights: effective remedies for third-country nationals in the Schengen Information System*. Brill. 199.

13 Kosta, E. (2013). *Consent in European data protection law* (Vol. 3). Martinus Nijhoff Publishers. *See also*, Custers, B., et al. (2022). *The role of consent in an algorithmic society – Its evolution, scope, failings and re-conceptualization*, in Kosta, E., Leenes, R., & Kamara, I. (Eds.). (2022). *Research handbook on EU data protection law*. Edward Elgar Publishing.

14 Commission of the European Communities, Proposal for a Council directive concerning the protection of individuals in relation to the processing of personal data, COM (90) 314 final – SYN 287, [1990] OJ C277/3, Article 12. For a detailed analysis on the evolution of the provisions of the Data Protection Directive on consent, see Kosta, E. (2013). *Consent in European data protection law* (Vol. 3). Martinus Nijhoff Publishers, 83.

15 Kosta, E. (2013). *Consent in European data protection law* (Vol. 3). Martinus Nijhoff Publishers, 95.

16 Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. O.J. (L 281/31) 24.11.1995.

17 Article 7 GDPR.

18 Article 8 GDPR.

19 European Union Charter of Fundamental Rights (CFR), Oct. 26, 2012, O.J. (C 83) 26.11.2012.

data protection as a fundamental right in the European legal order with special reference to consent, recognising that personal data 'must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law'.[20] In light of this enshrinement in the Charter, the right to data protection in general and consent in particular have been seen through a prism of rights-based individual autonomy.[21]

In searching for a 'normative justification for the broad scope of EU data protection law'[22] Lynskey has argued that 'individual control should be explicitly recognized as a facet of the right to data protection' because it can fulfil an instrumental function whereby it is 'one mechanism, amongst several, which seeks to render EU data protection law more effective'.[23] The way in which this can be achieved is through the reinforcement of individual rights over personal data and the ability to give and revoke consent if the controller relies on it. Individual control, according to Lynskey, should also 'guide the interpretation and application of the data protection rules to favour the individual'.[24] The European Data Protection Board (EDPB), the successor of the Article 29 Working Party, is responsible for the consistent application and enforcement of EU data protection law. As such, it issues binding decisions on disputes. It also issues guidance. Although not binding, it seeks to clarify the interpretation of data protection law provisions. According to the EDPB:

> if obtained in full compliance with the GDPR, consent is a tool that gives data subjects control over whether or not personal data concerning them will be processed. If not, the data subject's control becomes illusory, and consent will be an invalid basis for processing.[25]

The GDPR makes it clear that 'natural persons should have control of their own personal data'.[26] The notion of individual control has also featured, although not explicitly, in CJEU judgments. For instance, in a recent judgment, the Court highlighted that the right to access obliges data controllers to inform data subjects about the identities of recipients

20 Article 8(2) CFR.

21 Kosta, E. (2013). *Consent in European data protection law* (Vol. 3). Martinus Nijhoff Publishers. *See also* De Hert, P., & Gutwirth, S. (2009). Data protection in the case law of Strasbourg and Luxemburg: Constitutionalisation in action. In *Reinventing data protection?* (pp. 3-44). Dordrecht: Springer Netherlands

22 Lynskey, O. (2015). *The foundations of EU data protection law*. Oxford University Press.

23 Ibid. 229.

24 Ibid. 258.

25 European Data Protection Board, *Guidelines 05/2020 on consent under Regulation 2016/679 Version 1.1* (4 May 2020), 5.

26 Recital 7 GDPR. *See also* Recital 68 GDPR in relation to automated decision-making.

of their personal data, as this right 'must enable the data subject to verify not only that the data concerning him or her are correct, but also that they are processed in a lawful manner'.[27]As such, the Court seems to agree with the referring court that interpreting the right of access otherwise would 'seriously undermine the effectiveness of the legal remedies available to the data subject for the protection of his or her data'.[28] As a result, this chapter adopts the framing of consent as being tied to individual control and empowerment, though other scholars have considered the right to data protection from other, non-rights based perspectives.[29]

This chapter also focuses on transparency and its requirements set out in Articles 12 to 14 GDPR. This is because 'transparency requirements are not only an additional and separate obligation … but also an indispensable and constitutive part of the legal basis'.[30] To illustrate, in the case of the consent legal basis, it is stipulated that personal data can be processed if data subjects give consent 'to the processing of his or her personal data for one or more specific purposes'.[31] This implies that the data subject must be aware and understand what personal data are to be processed and for what purposes. Were it not for this information about the processing, data subjects would be unable to decide if they wished to give or withdraw consent because they would not know what they were consenting to or withdrawing consent from. Therefore, transparency on processing is an inextricable part of the legal basis. At the same time, transparency is also an additional set of obligations, which can be seen to play a supportive role in helping data subjects exercise control over their personal data, including in situations where the controller relies on the consent legal basis. According to the Article 29 Working Party:

> the data subject should be able to determine in advance what the scope and consequences of the processing entails and that they should not be taken by surprise at a later point about the ways in which their personal data has been used.[32]

27 Case C-154/21, *RW v Österreichische Post AG*, 12.01.2023, ECLI:EU:C:2023:3 para 32.

28 Ibid. 24.

29 *See, e.g.* Mahieu, R. L. P. (2023). The right of access to personal data in the EU: a legal and empirical analysis, 322: from the perspective of the right to access, the author argues that the right's justification lies less in individual control and more in 'equality, due process, and accountability of power'. *See also,* Gellert, R. (2020). *The risk-based approach to data protection*. Oxford University Press. The author argues that the GDPR is an example of a risk-based approach to regulation, rather than a rights-based one.

30 European Data Protection Board, *Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR)*, para 126.

31 Article 6(1)(a) GDPR.

32 Article 29 Working Party, *Guidelines on transparency under Regulation 2016/679* (11 April 2018), 7.

Thus, the purpose of transparency is to:

> empower data subjects to hold data controllers and processors accountable and to exercise control over their personal data by, for example, providing or withdrawing informed consent and actioning their data subject rights.[33]

As such, transparency requirements are framed as playing a supporting role in helping data subjects exercise control over their personal data.

## 2.2 The Requirements for Valid Consent and Transparent Data Processing

### 2.2.1 Valid Consent

Before assessing whether the personal data processing of three very large platforms is likely to meet some of the main requirements for valid consent and transparency as outlined in the GDPR, it is worth elucidating what these requirements are and how they have been interpreted. In Article 4(11) GDPR, consent is defined as:

> any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to him or her.

The burden is on the data controller who determines the means and purposes of the processing to demonstrate they have obtained valid consent.[34] *Freely given* consent means that the user should have a real choice to decide if they wish to agree to the processing of their personal data[35] and that the data subject should have the right to withdraw their consent at any time.[36] Consent is not freely given if there is a clear imbalance between the data subject and the data controller.[37] In *Meta Platforms*, the CJEU Court ruled that the dominance of a company in a relevant market can be a factor

> in assessing whether the user of that [company] has validly and, in particular, freely given consent, since that circumstance is liable to affect the freedom of choice of

33 Ibid. 5.
34 Article 7(1) GDPR.
35 European Data Protection Board, *Guidelines 05/2020 on consent under Regulation 2016/679 version 1.1.* (4 May 2020), 7.
36 Article 7(3) GDPR; European Data Protection Board (EDPB), *Guidelines 05/2020 on consent under Regulation 2016/679 version 1.1.* (4 May 2020), 7.
37 Article 43 GDPR.

> that user, who might be unable to refuse or withdraw consent without detriment, as stated in recital 42 of the GDPR.[38]

This is particularly relevant to very large platforms that are found to be dominant. Their practices could be shaken by the need to prove that consent was freely given despite the power imbalances inherent in their relationship to data subjects. However, it remains to be seen if this requirement challenges the power structures and imbalances in the online environment in practice and whether national data protection authorities will rely on the dominant position of the data controller in a relevant market when examining the validity of user consent.

To be valid, consent also needs to be specific, which according to the EDPB 'aims to ensure a degree of user control and transparency for the data subject'.[39] It is also closely linked to the requirement of informed consent.[40] In order to comply with the specificity requirements, the EDPB requires that:

> the data controller must apply i. Purpose specification as a safeguard against function creep, ii. Granularity in consent requests [which relates to the freely given requirement], and iii. Clear separation of information related to obtaining consent for data processing activities from information about other matters.[41]

Valid consent also needs to be *informed*, which closely relates to the transparency obligations discussed below. The EDPB highlighted that: 'depending on the circumstances and context of a case, more information may be needed to allow the data subject to genuinely understand the processing operations at hand'.[42] Finally, the GDPR added the requirement that consent must be *unambiguous*: it must be expressed by a statement or a clear affirmative action. The CJEU clarified that 'only active behaviour on the part of the data subject with a view to giving his or her consent may fulfil that requirement'.[43]

### 2.2.2 Consent for the Processing of Sensitive Categories of Personal Data

The processing of special categories of personal data is prohibited by Article 9(1) GDPR in principle. However, one of the exceptions to the prohibition is if 'the data subject has

38 Case C-252/21 *Meta Platforms Inc. and Others v Bundeskartellamt*, Judgement of the Court of Justice of the European Union (Grand Chamber) of 4 July 2023 (ECLI:EU:C:2023:537), para 148.

39 European Data Protection Board (EDPB), *Guidelines 05/2020 on consent under Regulation 2016/679 version 1.1.* (4 May 2020), 13.

40 Ibid. 14.

41 Ibid. 14.

42 Ibid. 16.

43 Case C-673/17 *Bundesverband der Verbraucherzentralen und Verbraucherverbände – Verbraucherzentrale Bundesverband v Planet49 GmbH*, 1.10.2019 ECLI:EU:C:2019:801, para. 54.

given explicit consent' to the processing 'for one or more specified purposes'.[44] The processing of special categories is regulated more strictly because 'their processing could create significant risks to the fundamental rights and freedoms'[45] of the data subject, for instance the risk of discrimination.[46] Special categories of personal data are defined in Article 9(1) GDPR as data:

> revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation.

The approach taken by the GDPR has been described in the literature as 'contextual', where special categories of personal data are processed whenever 'given the overall context, it would be possible to draw a conclusion from the data that might be sensitive in nature',[47] rather than focusing on the intent of the data controller to 'draw conclusions from the processing of particular data that could be regarded as being sensitive in nature'.[48] The contextual approach has been cemented by the CJEU in a recent case where it ruled that:

> the publication … on [a] website … of personal data that are liable to disclose *indirectly* the sexual orientation of a natural person constitutes processing of special categories of personal data.[49]

Therefore, what constitutes the processing of special categories of personal data has been interpreted very broadly.[50]

Moreover, in *Meta Platforms*, the CJEU ruled that individuals' visits to websites or apps and any data entered therein that reveal categories of personal data that are deemed

44 Article 9(2)(a) GDPR.

45 Recital 51 GDPR.

46 See for a discussion of discrimination as a harm resulting from data processing e.g. Lynskey, O. (2015). *The foundations of EU data protection law.* Oxford University Press.

47 Quinn, P., & Malgieri, G. (2021). The difficulty of defining sensitive data—The concept of sensitive data in the EU data protection framework. *German Law Journal*, *22*(8), 1583-1612. For an example of this approach *see e.g.* Datatilsynet, *Administrative fine – Grindr LLC* (2021), <https://www.datatilsynet.no/contentassets/8ad827efefcb489ab1c7ba129609edb5/administrative-fine---grindr-llc.pdf>

48 Ibid. 1591.

49 Case C-184/20 *OT v Vyriausioji Tarnybinės Etikos Komisija*, 1.09.2022 ECLI:EU:C:2022:601, para. 128.

50 For a discussion of the implications of such a broad scope *see* Quinn, P., & Malgieri, G. (2021). *The difficulty of defining sensitive data—The concept of sensitive data in the EU data protection framework*. German Law Journal, 22(8), 1583-1612.

sensitive trigger the application of Article 9(1) GDPR.[51] The Court also ruled that the same is true 'where a set of data containing both sensitive data and non-sensitive data is … collected *en bloc* without it being possible to separate the data items from each other'. [52] In other words, where there is at least one sensitive data item, this triggers the application of Article 9(1) GDPR. This is particularly relevant to online platforms. Owing to the automated nature of a lot of the data collection on the internet, it seems that, in the online context, having situations where sensitive and non-sensitive data are not collected '*en bloc*' would be rare. The consequence would be that Article 9(1) GDPR would apply to these processing operations. If special categories are deemed to be processed, and consent is the legal basis justifying such processing, the conditions for valid consent in the GDPR apply, and in addition to those, Article 9(2)(a) GDPR stipulates that consent should be 'explicit'. This requirement has been interpreted to mean that the controller must obtain an 'express statement of consent'. In an online context, this could be given by 'filling in an electronic form, by sending an email, by uploading a scanned document carrying the signature of the data subject, or by using an electronic signature', although the explicit consent need not necessarily be provided in a written form.[53]

### 2.2.3 Parental Consent for the Processing of Children's Personal Data in the Context of Information Society Services

Article 8 of the GDPR contains a parental consent requirement before the offering of 'information society services' directly to children under the age of sixteen (unless a lower national age threshold between thirteen and sixteen applies). The US Children's Online Privacy Protection Act (COPPA)[54] is one of the first pieces of legislation adopted to specifically protect the privacy of minors under thirteen years of age online. Although many European Member States considered minors ranging from fourteen to sixteen years to be competent to consent to the processing of their data,[55] the GDPR allows Member States to lower the bar for required parental consent to the age of thirteen, influenced by COPPA.[56] Thus, the processing of children's personal data as part of an online service, such as Xbox Games, insofar as these are offered directly to children, falls within the scope of Article 8 GDPR. As such, the processing can be legitimised 'if and to the extent

51 Case C-252/21 *Meta Platforms Inc. and Others v Bundeskartellamt*, Judgement of the Court of Justice of the European Union (Grand Chamber) of 4 July 2023 (ECLI:EU:C:2023:537), para 73.

52 Ibid. para 89.

53 European Data Protection Board, *Guidelines 05/2020 on consent under Regulation 2016/679 version 1.1.* (4 May 2020), 20-21.

54 Children's Online Privacy Protection Act 1998, 15 U.S.C. 6501–6505.

55 Dowty, T., & Korff, D. (2009). *Protecting the virtual child: the law and children's consent to sharing personal data*. National Youth Agency for Action on Rights for Children.

56 Macenaite, M., & Kosta, E. (2017). Consent for processing children's personal data in the EU: Following in US footsteps?. *Information & Communications Technology Law*, *26*(2), 146-197.

that consent is given or authorised by the holder of parental responsibility over the child'.[57] To comply with Article 8 GDPR, controllers are expected to 'make reasonable efforts to verify' whether the data subject is a child, and 'that consent is given or authorised by the holder of parental responsibility over the child, taking into consideration available technology'.[58]

### 2.2.4 Consent for Cookies

The ePrivacy Directive applies when the processing of personal data takes place via cookies. Article 5(3) ePrivacy Directive requires that the storing of information on the terminal equipment of users or subscribers or the gaining of access to such information is only allowed when the user or the subscriber has been provided with clear and comprehensive information and has given their consent.[59] Article 5(3) ePrivacy Directive has a broad scope and applies to any entity that stores information on a user device or accesses such information, in deviation from the general scope of application of the ePrivacy Directive.[60] In its judgment in *Planet 49*, the CJEU confirmed that valid consent should have the same meaning as in the GDPR and must be the result of an active behaviour of the user.[61] It may seem strange that Article 5(3) of the ePrivacy Directive contains a specific information requirement (provision of clear and comprehensive information). This would at first seem to overlap with the information and transparency requirements under the GDPR. However, the scope of Article 5(3) ePrivacy Directive covers not only personal data, but any kind of information. Accordingly, the inclusion of the information requirement in Article 5(3) ePrivacy Directive renders such provision mandatory, even if no processing of personal data takes place.[62] The Article 29 Working Party, after having examined a wide range of consent mechanisms that have been

57 Ibid.

58 Article 8(2) GDPR.

59 The ePrivacy Directive was amended in 2009 by the Citizens' Rights Directive: Directive 2009/136/EC of the European Parliament and the Council of the European Union amending Directive 2002/22/EC on universal service and users' rights relating to electronic communications networks and services, Directive 2002/58/EC concerning the processing of personal data and the protection of privacy in the electronic communications sector and Regulation (EC) No 2006/2004 on cooperation between national authorities responsible for the enforcement of consumer protection laws ('Citizens' Rights Directive') O.J. (L 337/11) 18.12.2009. In January 2017 the European Commission published a proposal for a Regulation concerning the respect for private life and the protection of personal data in electronic communications and repealing Directive 2002/58/EC (Regulation on Privacy and Electronic Communications), <http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=41241> which has not been adopted yet.

60 Kosta, E. (2013). *Consent in European data protection law* (Vol. 3). Martinus Nijhoff Publishers, 261;293.

61 Case C-673/17 *Bundesverband der Verbraucherzentralen und Verbraucherverbände – Verbraucherzentrale Bundesverband v Planet49 GmbH*, 1.10.2019 ECLI:EU:C:2019:801, paras 70-71.

62 Debusseré, F. (2005). The EU e-privacy directive: a monstrous attempt to starve the cookie monster?. *International journal of law and information technology*, *13*(1), 70-97. Kosta, E. (2013). Peeking into

deployed by website operators across Europe, published a working document that provides guidance on how valid consent could be obtained for cookies.[63] It stressed that users should be provided with specific information, including for processing purposes, an indication of cookies from third parties and retention periods.[64] The information should be provided 'before cookies are set or read'.[65] Lastly, 'the consent mechanism should present the user with a real and meaningful choice regarding cookies', and the user must be able to choose what cookies they wish to consent to with 'granularity'.[66]

### 2.2.5 Transparent Data Processing

The GDPR requires that, for personal data processing to be transparent, information about the processing must be provided to the data subject in writing or 'by other means, including where appropriate, by electronic means'.[67] In the online context, the most common mode of information provision is via privacy and cookie policies/statements, such as the ones analysed in this chapter. Article 12(1) GDPR requires that information provided to data subjects be in 'concise, transparent, intelligible and easily accessible form, using clear and plain language'. This has been interpreted by the Article 29 Working Party to mean that 'data controllers should present the information/ communication efficiently and succinctly in order to avoid information fatigue', that information 'should be understood by an average member of the intended audience' in order to be 'intelligible', and that '[the] data subject should not have to seek out the information'.[68] Data controllers 'should also separately spell out in unambiguous language what the most important consequences of the processing will be'.[69] The 'clear and plain language' requirement in Article 12(1) GDPR has been interpreted to mean that information:

> should be provided in as simple a manner as possible, avoiding complex sentence and language structures. The information should be concrete and definitive; it should not be phrased in abstract or ambivalent terms.... In particular the purposes of, and legal basis for, processing the personal data should be clear.[70]

the cookie jar: the European approach towards the regulation of cookies. *International journal of law and information technology*, *21*(4), 380-406.

63 Article 29 Working Party, *Working document 02/2013 providing guidance on obtaining consent for cookies (WP208)* (2013)

64 Ibid. 3.

65 Ibid. 4.

66 Ibid. 5.

67 Article 12(1) GDPR

68 Article 29 Working Party, *Guidelines on transparency under Regulation 2016/679* (11 April 2018), 7-8.

69 Ibid. 7.

70 Ibid. 8-9.

Words such as 'may' and 'might' should therefore be avoided because they do not describe data processing in a sufficiently clear manner.[71]

Articles 13 and 14 GDPR set out the categories of information that controllers must provide when data are collected directly from data subjects (Article 13 GDPR) and when data are obtained from other sources (Article 14 GDPR). If the controller collects the personal data from the data subject, the information that must be provided 'at the time when personal data are obtained'[72] include, inter alia, 'the purposes of the processing for which the personal data are intended as well as the legal basis for the processing', and 'the recipients or categories of recipients of the personal data, if any'.[73] Articles 13 and 14 GDPR also foresee that the controller might need to provide additional information to 'ensure fair and transparent processing', for instance 'the period for which the personal data will be stored', the existence of the right to withdraw consent at any time and so forth.[74]

## 3. Very Large Platforms' Processing of Personal Data on the Basis of Consent – Can the Transparency and Consent Requirements of EU Data Protection Law Be Met?

This part of the chapter is dedicated to analysing excerpts of privacy policies from three very large platforms – namely Google, Meta and Microsoft – against some of the main requirements for transparency and valid consent. This is done to ascertain the extent to which these could be met. This is done so we can reflect on whether consent is fit for the purpose of empowering individuals to exercise control over their personal data in section 5. Privacy policies, besides being a logical extension of the fairness and lawfulness principle, are a means to counter the information asymmetries between very large platforms and their users. They also facilitate the exercise of data subject rights, and they contribute to the accountability obligations of the platform providers.[75]

The following analysis is based on empirical evidence from the first author's PhD research on how five very large platforms[76] describe the processing of individuals' personal data. As part of this research, the privacy policies and other relevant web pages that describe the platforms' personal data processing (thirty-nine documents and web pages in total, excluding Terms of Service) were subjected to a qualitative content anal-

71 Ibid. 9.
72 Article 13(1) GDPR.
73 Article 13(1)(c) and Article 13(1)(d) GDPR.
74 Article 13(2) and Article 14(2) GDPR.
75 Van Alsenoy, B., Kosta, E., & Dumortier, J. (2014). Privacy notices versus informational self-determination: Minding the gap. *International Review of Law, Computers & Technology*, *28*(2), 185-203.
76 Namely Google, Amazon, Meta, Apple, Microsoft (GAMAM).

ysis to ascertain what categories of personal data the platforms process, the purposes for which they are processed and the legal bases that are used to justify the processing. The sample of platforms were chosen due to the scale at which they process online personal data, their size, reach and the 'systemic societal risks' they have been deemed to pose by the EU regulator.[77] For this chapter, the sample of documents mentioned above were analysed to ascertain how the platforms describe the processing of personal data based on consent. A few examples were then selected to be included and discussed in this chapter because they illustrate some of the broader issues related to meaningful transparency and consent in the online context. The chosen examples can also be assessed in more detail in terms of their compatibility with other tenets of EU data protection law, such as data minimisation and purpose limitation. However, the discussion here is mainly limited to consent and transparency, and even then, this chapter does not engage in a comprehensive analysis of all consent and transparency obligations. Finally, it should be noted that, while an effort was made to ensure that the examples are taken from the latest available versions of the privacy policies, by the time of publication the platforms may have changed the description of their processing activities, or even the legal ground for processing, depending on the circumstances.[78] Sections 3.1, 3.2 and 3.3 contain selected excerpts from the privacy policies of Google, Meta, and Microsoft respectively, which describe the processing of personal data based on consent. Section 3.4 discusses some of the main transparency issues. Finally, section 3.5 discusses some of the main issues with valid consent.

## 3.1 Processing Personal Data for the Purpose of Targeted Advertising, Including Cookies – Google

Targeted advertising entails processing personal data to target individuals with ads that they are likely to engage with, based on their traits (inferred or otherwise), interests and other data such as location, behavioural data and demographic data. Targeting therefore requires the processing of large volumes of personal data so it is able to draw inferences and make predictions about individuals' likelihood to engage with certain types of ads.

77 European Commission Staff Working Document, Impact Assessment Part 1 v.2 for Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (15 December 2020), <https://digital-strategy.ec.europa.eu/en/library/impact-assessment-digital-services-act>.

78 The legal ground can be changed, depending on the processing activity, from Article 6(1)(a) GDPR to another legal ground from Article 6(1). This is the case for the processing of personal data for the purpose of targeted advertising in particular, as the legal ground cannot be changed in the case of the setting of cookies, the processing of children's personal data in the context of information society services, or the processing of special categories of personal data.

**Google Privacy Policy**[79]

Depending on your settings, we may also show you personalized ads based on your interests.

We don't show you personalized ads based on sensitive categories, such as race, religion, sexual orientation, or health.

**Google: 'How Google Uses Information From Sites Or Apps That Use Our Services'**[80]

Ad Settings helps you control ads you see on Google services (such as Google Search or YouTube), or on non-Google websites and apps that use Google ad services. You can also LEARN HOW ads are personalized, opt out of ad personalization, and block specific advertisers.

Many websites and apps use Google services to improve their content and keep it free. When they integrate our services, these sites and apps share information with Google. For example, when you visit a website that uses advertising services like AdSense, including analytics tools like Google Analytics, or embeds video content from YouTube, your web browser automatically sends certain information to Google. This includes the URL of the page you're visiting and your IP address. We may also set cookies on your browser or read cookies that are already there. Apps that use Google advertising services also share information with Google, such as the name of the app and a unique identifier for advertising. Google uses the information shared by sites and apps to deliver our services, maintain and improve them, develop new services, measure the effectiveness of advertising, protect against fraud and abuse, and personalize content and ads you see on Google and on our partners' sites and apps.

**Google: 'How Google Uses Cookies'**[81]

Google uses cookies for advertising, including serving and rendering ads, [and] personalizing ads (depending on your settings)

79 *Google Privacy Policy* (15 December 2022), <https://policies.google.com/privacy?hl=en-US>.

80 *How Google Uses Information from Sites or Apps that Use our Services*, <https://policies.google.com/technologies/partner-sites?hl=en-US&hl=en_US>.

81 *How Google Uses Cookies*, <https://policies.google.com/technologies/cookies?hl=en-US>. *See also, Our Advertising and Measurement Cookies*, <https://business.safety.google/adscookies/?hl=en_US>.

### 3.2 Processing Special Categories of Personal Data – Meta

**Meta**[82]
Processing information with special protections that you provide so we can share it with those you choose, to provide, personalise and improve our products and to undertake analytics.

Information categories we use …. The actual information we use depends on your factual circumstances, but could include any of the following:
Your activity and information that you provide:
Any information with special protections that you choose to provide in your profile fields (such as your religious views, political views or who you are 'interested in'), or as part of surveys that you choose to participate in, and where you have given your explicit consent.

### 3.3 Processing Children's Personal Data – Microsoft

**Microsoft Xbox**[83]
You as the parent or guardian are required to consent to the collection of personal data from a child under 13 years old. With your permission, your child can have an Xbox profile and use the online Xbox network.

We collect information about your child's use of Xbox services, games, apps, and devices including:
Which games they play and apps they use, their game progress, achievements, play time per game, and other play statistics.
Content they add, upload, or share through the Xbox network, including text, pictures, and video they capture in games and apps.
Social activity, including chat data and interactions with other gamers, and connections they make (friends they add and people who follow them) on the Xbox network.

Microsoft uses the data we collect to improve gaming products and experiences—making it safer and more fun over time. Data we collect also enables us to provide your child with personalised, curated experiences. This includes connecting them to games, content, services, and recommendations.

82 *Meta Privacy Policy* (5 April 2023), <https://mbasic.facebook.com/privacy/policy/printable/?_rdr/>.
83 *Microsoft Privacy Policy* (February 2023), <https://privacy.microsoft.com/en-us/privacystatement>.

> In order to help make the Xbox network a safe gaming environment and enforce the Community Standards for Xbox, we may collect and review voice, text, images, videos and in-game content (such as game clips your child uploads, conversations they have, and things they post in clubs and games).

## 3.4 Transparency Issues

### 3.4.1 Categories of Personal Data

As outlined in section 2.2.5, data controllers must inform individuals about the categories of personal data they process. In all of the excerpts above, the categories are described in vague terms that make it difficult to ascertain what personal data are processed. This subsection focuses on Google and Meta in particular. For the processing of personal data for targeted advertising, Google processes data based on 'interests'. These must be inferred by the platform, since the first section of the privacy policy mentions that it collects three types of personal data. These are: the information that data subjects provide, such as payment information; data subjects' activities on the services, such as terms searched and videos watched; and location data.[84] Which of these data are specifically used to make inferences about 'interests' is not clear, nor is it clear how processing the data is necessary for the purpose of targeted advertising. The policy mentions some categories of personal data are *not* processed for the purpose of targeting data subjects with ads, like special categories of personal data. To the authors, mentioning only interests without information on the categories of personal data that have led to their inference and only mentioning the categories of personal data that are not processed instead of those that are processed, is inconsistent with the requirement for 'clear and plain language' laid down in Article 12(1) GDPR, and with the requirement to provide the categories of personal data processed laid down in Article 13 GDPR.

It is also doubtful that Google does not process sensitive categories of personal data. This is first because, as shown in section 2.2.2, what constitutes the processing of sensitive categories is construed very broadly. The CJEU's recent finding, 'where a set of data containing both sensitive data and non-sensitive data is … collected *en bloc* without it being possible to separate the data items from each other', is particularly relevant here. In other words, if there is at least one sensitive data item, then Article 9(1) GDPR applies.[85] Accordingly, if the information provided by data subjects, their activity, and their location data reveal special categories *and* are collected 'en bloc' without the possibility of

84 *Google Privacy Policy* (15 December 2022), <https://policies.google.com/privacy?hl=en-US>.

85 Case C-252/21 *Meta Platforms Inc. and Others v Bundeskartellamt*, Judgement of the Court of Justice of the European Union (Grand Chamber) of 4 July 2023 (ECLI:EU:C:2023:537), para 89.

separating data items from each other, Google's processing would be prohibited in principle under Article 9(1) GDPR, unless it sought and obtained explicit consent under Article 9(1)(a) GDPR or relied on any other derogation in Article 9(2) GDPR. Second, Google's own settings in My Ad Center allow individuals to 'limit ads about sensitive topics, like weight loss and alcohol'.[86] This raises the question of how ads can be effectively targeted in terms of 'sensitive topics' such as weight loss and alcohol, without any inferences being made about the data subject's health status and conditions, which would constitute the processing of special categories of personal data. To be clear, it is not that ads could not be targeted on these topics without special categories of personal data being inferred, but that these inferences would most likely make the ads more effective. In other words, they would increase the likelihood of data subjects engaging with them, which is crucial for Google's bottom line.[87]

*Meta* is similarly unclear about the special categories of personal data it processes or how inferences are made about individuals. The policy mentions that two categories of personal data are processed: 'your activity' and 'information that you provide'. The policy explains that activity refers to 'all of the things you can do on our products'. Comparatively, the 'information that you provide' category, because of the use of 'such as', is non-exhaustive. It can include content, including posts, comments, audio, purchases and transaction data, '[t]he time, frequency, and duration of your activities', hashtags used and more. The use of non-definitive language makes it hard to ascertain whether other types of data would count as 'information that you provide'. Likewise, the category 'your activity' does not allow the data subject to understand exactly what personal data are processed. For instance, is it that only the activity on Meta products is processed to infer special categories of personal data? Does the category also include activity on non-Meta products, which is collected via/received from third parties 'whether or not you're logged in or have an account on our products'? Ultimately, given how broad the two categories of 'information you provide' and 'your activity' are, it gives the impression that virtually *all* types of data collected by Meta are processed to infer special categories. This raises the question of whether the processing is in any way limited to what is strictly necessary. Several data protection authorities from various Member States and the EDPB have criticised the Irish Data Protection Authority for not examining the lawfulness of Meta's processing of special categories of personal data, and the latter has ordered the Irish authority to investigate further.[88] The Irish authority rejected this call to investigate, and filed an action for annulment before the CJEU, claiming

86 *Google My Ad Center,* <https://myadcenter.google.com/home>.

87 Google's business model relies on advertising, which has generated 147 billion dollars in revenues in 2020. *See* Wieringa, R. (2022) *Google's Business Model* (The Value Engineers 4 February 2022), <https://www.thevalueengineers.nl/googles-business-models/>.

88 European Data Protection Board, *Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR)*, paras 147-152.

that the EDPB 'exceeded its competence under Article 65(1)(a) [GDPR]'.[89] It remains to be seen if the CJEU will agree. For now, the Irish authority's reluctance as Lead Supervisory Authority to investigate the lawfulness of Meta Ireland's processing of special categories of personal data does not provide much hope that the issues related to Meta discussed in this chapter will be addressed anytime soon, despite the high risks such processing poses for the rights and interests of data subjects.

### 3.4.2 Purposes for Processing

Per the transparency obligations outlined in section 2.2.5, data controllers must also inform data subjects about the purposes for which they process personal data. In all of the excerpts, the purposes are described in vague terms that make it difficult to determine what they entail; they are not sufficiently 'specified' within the meaning of the purpose limitation principle in Article 5(1)(b) GDPR. It is also difficult to ascertain how the categories of personal data outlined are limited to what is necessary for their fulfilment, as required by the data minimisation principle in Article 5(1)(c) GDPR. In the Google excerpt, there seem to be at least three processing operations described. The first is the processing of personal data (excluding cookies) for the purpose of targeted advertising. The second is the processing of personal data via cookies for the purpose of showing targeted advertising. Finally, there is the processing of data (including that from cookies) received from third parties who use Google advertising services, which are processed for several purposes, including targeted advertising and the delivery of the services. What targeted advertising actually entails is not explained, nor does Google indicate what the possible consequences of processing personal data for targeted advertising, including from cookies, for the fundamental rights and freedoms of data subjects could be. This means individuals are unlikely to understand the consequences of consenting to such processing just from reading the provided information. The same can be said of the other purposes mentioned, such as the 'delivery of services.'

The Meta policy claims that special categories of personal data are processed to 'provide, personalise, and improve our products and to undertake analytics'. In light of the purpose limitation principle and the Article 29 Working Party guidelines, these purposes are not described in a clear enough manner for the data subject to understand how their personal data are used. For instance, it is claimed that data are processed for the 'personalisation' of products, a word which Meta uses to refer to both targeted advertising and to personalising the news feed of users. However, it is not clear whether the personal data are processed for only one or both of these purposes, nor how the data themselves are needed for any of the purposes. The same can be said of the other pur-

89 Case T-84/23, *Data Protection Commission v European Data Protection Board* (11 April 2023), <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:62023TN0084>.

poses mentioned, such as in the context of the purpose of 'provid[ing] and improv[ing] our products'. It is not clear what the products referred to are (i.e. does this refer to all the Meta products described in the privacy policy or just some of them?) nor how the processing of special categories of personal data will help to provide or improve these products.

Likewise, in Microsoft's policy, the purposes for processing are not described in a clear enough manner for the data subject to understand how their personal data are processed. It should be stressed that when the data controller is targeting children, these 'do not lose their rights as data subjects to transparency simply because consent has been given/authorised by the holder of parental responsibility in a situation to which Article 8 of the GDPR applies'.[90] This means that the information about how personal data are processed should, even when Article 8 GDPR applies, be aimed at the child, and so 'any information… should be conveyed in clear and plain language or in a medium that children can easily understand'.[91] Microsoft appears to have made little effort to make the policy understandable to children. The additional 'Xbox data collection for kids'[92] page adds images, but the text is very similar to that in the privacy policy. Microsoft also has a separate, 'Privacy for young people'[93] webpage, which seems to cover all processing of children's data not just that in the context of Xbox. However, this is not linked in the main privacy policy, and it is hard to ascertain if this page would give children a better understanding of how their personal data are processed. From the excerpts, it therefore seems that it would likely be unclear what the various purposes entail for both parents and children. For instance, what would the processing of personal data for the purpose of creating 'a safe gaming environment' entail, beyond the enforcement of the Community Standards for Xbox? How are the categories of personal data of 'voice, text, images, videos' limited to what is strictly necessary to achieve this purpose?

## 3.5 Consent Issues

### 3.5.1 Informed

As outlined in section 2.2.1, consent must be informed to be valid. However, as shown in the previous section, transparency issues can be identified in all three platforms' descriptions of their processing activities. These issues create serious doubt about

90 Article 29 Working Party, *Guidelines on transparency under Regulation 2016/679* (11 April 2018), 10.

91 Ibid. 10-11. Unless the child is very young or pre-literate, which is unlikely to be the case for most children playing Xbox.

92 *Microsoft Xbox data collection for kids*, <https://www.xbox.com/en-US/family-hub/user-data-collection-kids>.

93 *Microsoft Privacy for Young People* (March 2023), <https://privacy.microsoft.com/en-US/young-people>.

whether any of the platforms would be able to obtain informed consent from data subjects. None of the platforms make it clear what the consequences of the processing would be, so the provided information is arguably not enough for data subjects to 'enable them to make informed decisions' and 'understand what they are agreeing to'.[94]

To reiterate, for Google, the analysis above has highlighted a lack of clarity and specificity of information about the categories of personal data that are processed, and the purposes for processing. The excerpts above seem to indicate that when users have consented to targeted advertising, the personal data used to target ads can include *all* types of personal data collected, in principle, except for special categories and other data explicitly mentioned in the privacy policy as not being processed. Whether such categories are in fact excluded is impossible to verify, but, as discussed above, this can be doubted.

As for Meta, the above analysis has likewise highlighted a lack of clarity and specificity regarding the categories of personal data that are processed, and the purposes for processing. The excerpts indicate that *all* types of personal data collected are processed to infer special categories of personal data, which are then used for a variety of purposes.

Microsoft's descriptions of categories and purposes are similarly vague. In addition, it is doubtful that children or their parents would be able to understand how their data are processed and for what purposes, and the consequences of such processing.

### 3.5.2 Specific and Freely Given

Consent also needs to be specific and freely given to be valid. Here, these two elements are analysed together. It is argued that the consent obtained from data subjects is unlikely to meet the 'specific' requirement for any of the three platforms' processing operations. Moreover, there are some issues that suggest the obtained consent might not be freely given. It is not clear from its policy how Google obtains specific consent for the processing of personal data (excluding cookies) for the purpose of targeted advertising. As already pointed out, the purpose of 'targeted advertising' is not sufficiently specific and is very broad, meaning that there is a risk of 'gradual widening or blurring of purposes for which data is processed, after a data subject has agreed to the initial collection of the data'.[95] In My Ad Center, individuals are asked if they want to 'turn on personalisation', which might indicate this is how Google asks for consent. After turning on personalisation individuals are given more settings to choose from. These allow them to see the topics of ads targeted to them based on their recent user activity, including 'sensitive' ones like weight loss, alcohol, dating, gambling, pregnancy and

94 European Data Protection Board (EDPB), *Guidelines 05/2020 on consent under Regulation 2016/679 version 1.1.* (4 May 2020), 15.

95 Ibid. 13.

parenting. The idea seems to be that data subjects 'choose topics and brands you want to see more or fewer ads about'.[96] At first, it might appear that these are consent settings that allow users to control what data are used for targeted advertising at a granular level. Yet upon closer inspection and considering the language used, these seem to be settings that control not the types of personal data Google processes in terms of target ads, but what ads the data subjects *see*. In other words, it seems that even if a user chooses not to see ads related to weight loss, for example, there seems to be no indication that this would result in Google processing less or different personal data. Creating the impression that data subjects can exercise further control over their data via settings in My Ad Center seems counter to the transparency and specific consent requirements in data protection law. Google should allow for more granular consent given how broad the purpose of 'targeted advertising' is. Google's practices might also conflict with the principle of fairness under Article 5(1)(a) GDPR. In its recent binding decision on the Irish Supervisory Authority, the EDPB found that Meta breached the principle of fairness because, inter alia, its presentation of the legal basis was misleading and deceptive.[97] There is, therefore, a case to be made that misleading data subjects to think they have more granular control over what data are processed for targeted advertising is in breach of the principle of fairness.

As for Meta, because different kinds of data are processed for so many (vague) purposes, data subjects essentially consent to a 'bundle of processing activities'.[98] This casts doubt on whether the specific consent requirement is met, as that would require data subjects to be able to refuse consent for each specific purpose for which special categories of personal data are processed. However, this does not appear to be possible. This requirement for granularity in specific consent is also closely linked to the freely given requirement, whereby consent is not considered to be freely given where 'the process/procedure for obtaining consent does not allow data subjects to give separate consent for personal data processing operations'.[99] In the Meta Platforms case, the CJEU made it clear that the dominance of a firm can be a factor in assessing whether the users of that firm have provided freely given consent to data processing, as 'the existence of such a dominant position may create a clear imbalance, within the meaning of recital 43 of the GDPR, between the data subject and the controller'.[100] Based on this, it ruled that

96 *Google My Ad Center,* <https://myadcenter.google.com/home>.

97 European Data Protection Board, *Binding Decision 3/2022 on the dispute submitted by the Irish SA on Meta Platforms Ireland Limited and its Facebook service (Art. 65 GDPR),* para 230.

98 European Data Protection Board, *Guidelines 05/2020 on consent under Regulation 2016/679 version 1.1.* (4 May 2020), 12.

99 Ibid. 12.

100 Case C-252/21 *Meta Platforms Inc. and Others v Bundeskartellamt*, Judgement of the Court of Justice of the European Union (Grand Chamber) of 4 July 2023 (ECLI:EU:C:2023:537), paras 148-149.

users must have the possibility to refuse to consent separately to the processing of personal data from third-party sources without being deprived of the social network service Meta offers.[101] If the referring Court finds that users are not given this possibility, then 'the consent of those users … must be presumed not to be freely given'.[102] As it has been established as part of this case that Meta is a dominant firm on the social network market, it can be argued that it must likewise provide for the opportunity to refuse consent separately to the processing of special categories of personal data in relation to each purpose, without being deprived of the social network service. Otherwise, users have no choice but to consent to use the service, so their consent cannot be considered as freely given. The Court's finding that dominance can effect whether consent is freely given could have similar implications for other very large platforms, like Google and Microsoft, if these are found to be dominant.

In the case of Microsoft, it is difficult to ascertain how consent is obtained from the holder of parental responsibility, and how the platform engages in user verification in order to ensure that it meets the requirements of Article 8 GDPR. In addition, it is unclear whether granular consent is obtained for the processing of children's personal data for each of the purposes outlined in the privacy policy. Since Microsoft seems to base all of its processing of children's personal data on consent, there seem to be some issues in relation to the requirement that it be freely given. For instance, the excerpt describes the processing of personal data for the purpose of making the Xbox environment 'safe' and enforcing the Community Standards. However, if holders of parental responsibility are indeed asked to consent for processing the child's personal data for 'safety' purposes, could the child still use the Xbox service if consent were to be withdrawn? If the holder of parental responsibility cannot withdraw consent without detriment, for instance without the child being 'barred' from the service or having the service downgraded in some other way,[103] then the consent cannot be considered freely given. In this context, consent might be an inappropriate legal basis for processing. Recently, in the United States, the FTC fined Microsoft for illegally processing children's personal data in the context of their Xbox service without parental consent in the period 2015 to 2020, contrary to COPPA requirements.[104] This should raise alarm bells for data protec-

101 Ibid, para 150.

102 Ibid, para 151.

103 European Data Protection Board (EDPB), *Guidelines 05/2020 on consent under Regulation 2016/679 version 1.1.* (4 May 2020), 13.

104 Federal Trade Commission, *FTC Will Require Microsoft to Pay $20 million over Charges it Illegally Collected Personal Information from Children without Their Parents' Consent* (5 June 2023) <https://www.ftc.gov/news-events/news/press-releases/2023/06/ftc-will-require-microsoft-pay-20-million-over-charges-it-illegally-collected-personal-information>.

tion authorities who might need to investigate whether the same violations took place in relation to children's personal data in the EU.

### 3.5.3 Explicit Consent

Since Meta processes special categories of personal data, there is an additional requirement it must fulfil for consent to be valid: consent must be explicit. However, it is argued that Meta does not meet this requirement. As mentioned above, Meta's description of how special categories of personal data are processed seems to imply that whatever data subjects do on its products can and is used to infer special categories of personal data. Therefore, if a user were to use a hashtag commonly associated with an eating disorder to search Instagram or Facebook, for example, or if they were to tag one of their images or videos with such a hashtag, Meta would infer data concerning health through its analytics processes. However, data subjects are not asked to provide explicit consent before engaging in all activities on the Meta products, such as scrolling through the newsfeed and moving the cursor on the page, posting status updates, searching or using hashtags. On the contrary, the Meta privacy policy seems to imply that whenever its 411 million EU monthly active users[105] make use of its products, consent for the processing of special categories of personal data is implicit. Needless to say, implicit consent is unlawful under EU data protection law and does not even meet the prerequisite for 'regular' unambiguous consent, which should be 'a statement of clear affirmative action'.[106]

Therefore, it cannot possibly meet the higher threshold for 'explicit' consent for the processing of special categories of personal data, which would require that the data subject expresses their consent via, for instance, a written statement, filling in an electronic form, sending an email, etc.[107] On the other hand, assuming that explicit consent is obtained (which it appears it is not), how can data subjects withdraw it? For consent to be valid, data subjects should be informed about, and able to withdraw consent without detriment. From the privacy policy alone, and without an account on Meta services, it is difficult to ascertain if this is the case. Lastly, Meta is also a dominant firm on the market for social networks. The power imbalance vis-à-vis data subjects allows it to impose the processing of special categories of personal data as part of its service, and this data is then used in ways that data subjects cannot understand based on the vague descriptions of purposes. This practice is in serious conflict with the requirements that consent be freely given and explicit, among others. Regulatory action is necessary,

105 *Facebook monthly active users (MAU) in Europe as of the 1st quarter 2023*, Statista (9.05.2023), <https://www.statista.com/statistics/745400/facebook-europe-mau-by-quarter/>.

106 Article 4(11) GDPR.

107 European Data Protection Board (EDPB), *Guidelines 05/2020 on consent under Regulation 2016/679 version 1.1.* (4 May 2020), 20.

given the risks the processing poses for the rights and interests of data subjects, but it remains to be seen if and when this will materialise.

## 4. Consent and Transparency – Tools With a Disempowering Effect in the Online Context?

The previous section shows that three very large platforms' online processing of personal data falls short of meeting some of the main transparency and consent requirements set out in EU data protection law. The transparency issues identified make it hard to understand how data are used and for what purposes. For all three platforms, it seems that not all the information about the processing is provided in the privacy policies. Individuals must also often have an account on the service to receive more information about the processing, to give consent and to manage consent via settings. There is also evidence that the consent obtained is often not informed, specific, freely given or, in the case of the processing of special categories of personal data, explicit. This casts doubt on whether these processing activities can be lawfully based on consent, as there is a high risk that the consent obtained is illusory and unable to fulfil the purpose of empowering data subjects to exercise control over their personal data. The findings, therefore, are in stark contrast to the aims and rationale of consent and transparency in EU data protection law. As such, further investigation of these and other such practices is needed. However, the findings discussed here are consistent with academic literature, which has criticised consent in EU data protection law and its role in the online context over the years. We turn to this literature now, before concluding with our reflections on the role of consent.

At the time of the Data Protection Directive, one issue was raised related to the 'diverging ways in which consent has been implemented in various Member States', which was thought to 'pose threats to the uniform application of the relevant provisions across Europe'.[108] This lack of harmonisation also worked to hinder the realisation of consent's role of uniformly empowering individuals across the EU to exercise control over their personal data. However, legal scholars recognised that, even if the conditions for valid consent in data protection law were to be harmonised, which they eventually were in the GDPR, questions regarding whether consent could fulfil its empowering role in a rapidly changing technological landscape would remain.[109] In particular, scholars worried that the consent model, which assumes 'the data subject makes a free, conscious and rational

108 Kosta, E. (2013). *Consent in European data protection law* (Vol. 3). Martinus Nijhoff Publishers, 147.

109 *See e.g.* Koops, B. J. (2014). The trouble with European data protection law. *International data privacy law*, *4*(4), 250-261, arguing that the focus on informational self-determination is misguided. *See also*, Moerel, L., & Prins, C. (2016). Privacy for the homo digitalis: Proposal for a new regulatory framework for data protection in the light of Big Data and the internet of things. *Available at SSRN 2784123*.

decision to consent based on the available information' would 'not have the effect … desired by the legislator'[110] because, in practice in the online environment, there is an information overload, a consent overload and an absence of meaningful choice.[111] These criticisms are informed by empirical research on companies' practices and individuals' online behaviour, as well as by insights from behavioural economics. The latter have allowed legal scholars to argue that the concept of rational consent has limitations because individuals make decisions under a bounded rationality.[112] US scholars simultaneously also expressed serious doubts that consent as part of the 'privacy self-management' paradigm would be able to deliver on its promises of granting individuals control over their personal data.[113] Now, more than a decade later, the criticisms levelled at consent in the context of the online environment are still relevant. The issue of consent's role has become even more critical in light of technological advancements, particularly in data analytics and AI, which have improved and expanded on what controllers can do with personal data.[114]

Post-GDPR, the issue of consent overload or fatigue remains, as do the issues of information overload and 'take-it-or-leave-it' choices, whereby individuals have little choice to not consent if they wish to use certain services.[115] More recently, research on so-called 'dark patterns' has explored how user interfaces (e.g. in cookie banners) can manipulate individuals into consenting to the processing of their personal data.[116] Fur-

110 Schermer, B. W., Custers, B., & Van der Hof, S. (2014). The crisis of consent: How stronger legal protection may lead to weaker consent in data protection. *Ethics and Information Technology*, *16*(2), 176.

111 Ibid.

112 *See e.g.* Custers, B., van Der Hof, S., Schermer, B., Appleby-Arnold, S., & Brockdorff, N. (2013). Informed consent in social media use-the gap between user expectations and EU personal data protection law. *SCRIPTed*, *10*, 435. Carolan, E. (2016). The continuing problems with online consent under the EU's emerging data protection principles. *Computer Law & Security Review*, *32*(3), 462-473. Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE security & privacy*, *3*(1), 26-33. Monteleone, S. (2015). Addressing the failure of informed consent in online data protection: learning the lessons from behaviour-aware regulation. *Syracuse J. Int'l L. & Com.*, *43*, 69. Custers, B. et al, *Consent and Privacy*, In: Müller, A., & Schaber, P. (Eds.). (2018). *The Routledge handbook of the ethics of consent* (pp. 119-130). London, UK: Routledge.

113 *See e.g.* Solove, D. J. (2012). Introduction: Privacy self-management and the consent dilemma. *Harv. L. Rev.*, *126*, 1880. *See also* Hartzog, W., & Richards, N. (2020). Privacy's constitutional moment and the limits of data protection. *BCL Rev.*, *61*, 1687, for a more recent critical reflection on the individual control aspects of EU data protection law by US scholars.

114 For a technical discussion, *see e.g.* Shi, Y. (2022). Advances in big data analytics. *Adv Big Data Anal.*

115 *See e.g.* Zuiderveen Borgesius, F. J., Kruikemeier, S., Boerman, S. C., & Helberger, N. (2017). Tracking walls, take-it-or-leave-it choices, the GDPR, and the ePrivacy regulation. *Eur. Data Prot. L. Rev.*, *3*, 353. Van Alsenoy, B., Kosta, E., & Dumortier, J. (2014). Privacy notices versus informational self-determination: Minding the gap. *International Review of Law, Computers & Technology*, *28*(2), 185-203.

116 *See e.g.* Nouwens, M., Liccardi, I., Veale, M., Karger, D., & Kagal, L. (2020, April). Dark patterns after the GDPR: Scraping consent pop-ups and demonstrating their influence. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-13); Graßl, P. A. J., Schraffenberger, H. K., Zuiderveen Borgesius, F. J., & Buijzen, M. A. (2021). Dark and bright patterns in cookie consent

thermore, research has questioned the extent to which the requirement of parental consent can adequately protect children,[117] and whether the requirements for explicit consent for the processing of special categories of personal data are actually met.[118] Consent's ability to empower individuals to take control over their personal data depends on the broader context of the processing activity and how consent is obtained. For the online context, which is the focus of this chapter, literature has consistently shown that consent's ability to fulfil its empowering role envisioned by the EU legislator is severely limited in practice, due to both companies' practices in the online environment and individual decision-making. This means that, in practice, the consent obtained hardly meets all requirements for validity.[119] As such, because of this chapter's empirical findings and the outlined criticisms, it seems that consent in the online context tends to have a disempowering effect as opposed to an empowering one.

## 5. Concluding Thoughts: Reflecting on the Current and Future Role of Consent in EU Data Protection Law

This chapter sought to re-visit a debate that is at least as old as EU data protection law itself, namely that of whether consent can, in an ever more digitised world, fulfil its role in empowering data subjects to exercise control over their personal data. This section discusses two main approaches to preserving consent's empowering role. In doing so, it reviews some suggestions put forward in the literature, and we propose some of our own. However, determining exactly what should be changed and how remains a difficulty. A third and final approach is considered: that of questioning the very business model of personal data exploitation, which most market actors online rely on to various

requests. *Journal of Digital Social Research* 3.1; Luguri, J., & Strahilevitz, L. J. (2021). Shining a light on dark patterns. *Journal of Legal Analysis*, *13*(1), 43-109.

117 *See e.g.* Jasmontaite, L., & De Hert, P. (2015). The EU, children under 13 years, and parental consent: a human rights analysis of a new, age-based bright-line for the protection of children on the Internet. *International Data Privacy Law*, *5*(1), 20-33. Van der Hof, S. (2016). I agree, or do I: a rights-based analysis of the law on children's consent in the digital world. *Wis. Int'l LJ*, *34*, 409; van der Hof, S., Lievens, E., & Milkaite, I. (2019). The protection of children's personal data in a data-driven world: A closer look at the GDPR from a children's rights perspective. In *Monitoring children's rights in the Netherlands: 30 years of the UN convention on the rights of the child* (pp. 77-123). Leiden University Press; Macenaite, M., & Kosta, E. (2017). Consent for processing children's personal data in the EU: Following in US footsteps?. *Information & Communications Technology Law*, *26*(2), 146-197.

118 *See e.g. Mantovani, E., & Quinn, P. (2014). mHealth and data protection–the letter and the spirit of consent legal requirements. International Review of Law, Computers & Technology, 28(2), 222-236*; Pormeister, K. (2017). Informed consent to sensitive personal data processing for the performance of digital consumer contracts on the example of "23andMe". *Journal of European Consumer and Market Law*, *6*(1); Quinn, P., & Malgieri, G. (2021). The difficulty of defining sensitive data—The concept of sensitive data in the EU data protection framework. *German Law Journal*, *22*(8), 1583-1612.

119 Kosta, E. (2013). *Consent in European data protection law* (Vol. 3). Martinus Nijhoff Publishers.

degrees, and reflecting on whether individuals should be protected rather than empowered in the online context.

This chapter has shown, consistent with previous literature on consent, that there is a clear disconnect between the letter of the law, which positions consent as a tool for empowerment, and the reality of market actors' data processing practices. If we accept that the problem does not lie with consent itself and its validity requirements, then it might be concluded that this disconnect is solely due to market actors that undermine the ideal of consent through, inter alia, vague descriptions of processing, not allowing for specific or explicit consent and so forth. As a result, one possible way for consent to fulfil its empowering role is through tackling the practices of market actors that undermine it. This approach seeks to correct the market and align its operation with the requirements set out in the law for transparency and valid consent. To achieve this, we would need an empirically-informed understanding of market actors' practices. However, empirical research on privacy policies can only go so far in uncovering how personal data are processed, as we require an understanding of how data are processed de facto. Nevertheless, more empirical research and the filing of complaints to Data Protection Authorities could be a start.

The Achilles' heel of this approach is enforcement issues that have proved to be major stumbling blocks for the effectiveness of the EU data protection law regime. For instance, it is well known that data protection authorities in Member States are understaffed and underfunded.[120] In particular, when it comes to very large platforms, the one stop shop mechanism that often sees the Irish Data Protection Authority as Lead Authority has been criticised for its effectiveness in curbing unlawful data processing practices.[121] Consent itself has been criticised heavily in the literature. However, it is not always clear how market actors undermine its role through practices, such as the ones identified in section 3 of this chapter. The approach outlined so far would thus keep the legal standard of consent in EU data protection law as it is and shift the focus to the behaviour of market actors, on whom the onus is to comply with the law. There have been many calls from academia[122] and civil society[123] to strengthen enforcement. It also seems that some

120 *See e.g. Data protection: 80% of national authorities underfunded, EU bodies 'unable to fulfil legal duties'*, Statewatch (30 September 2022), <https://www.statewatch.org/news/2022/september/data-protection-80-of-national-authorities-underfunded-eu-bodies-unable-to-fulfil-legal-duties/>.

121 For a study of enforcement against big tech companies, see Irish Council for Civil Liberties, *5 years: GDPR's crisis point: ICCL report on EEA data protection authorities* (2023), <https://www.iccl.ie/wp-content/uploads/2023/05/5-years-GDPR-crisis.pdf>. *See also,* Miglio, A. *Facebook Ireland: The Scope of the 'One-Stop Shop' under the GDPR and how to react to enforcement bottlenecks*, EU Law Live (17 June 2021), <https://iris.unito.it/bitstream/2318/1836549/1/OpEd_AlbertoMiglio_17June2021.pdf>.

122 See, e.g. Lancieri, F. *Narrowing Data Protection's Enforcement Gap* (2022) Maine Law Review 74.1.

123 *See, e.g.* the work of NGO 'None of Your Business' (NOYB), <https://noyb.eu>. Also, the work of the Irish Council for Civil Liberties (ICCL), <https://www.iccl.ie/digital-data/>.

initiatives are underway, for instance the EU Commission's plan to streamline cooperation between Data Protection Authorities in cross-border cases.[124]

However, even prior to the GDPR, it was anticipated that this approach would not work in practice. Issues of enforcement have seen those concerns materialise, leading to the criticisms levelled at consent itself rather than (solely) at the practices of market actors. As discussed in section 4, critics worried that the conditions for valid consent could hardly be met in the online context in particular, not least because how consent applies to individual decision-making is flawed, as individuals do not always get to make fully-informed, rational decisions. As such, consent and its conditions for validity, as well as its ideal of empowerment, are seen to be misguided, flawed and unachievable. The practices of market actors in the online context also serve to expose this failure to translate how individuals behave and make decisions into a legal standard that represents that reality. The only way this approach would see consent as an empowering tool is if the legal standard of consent would be changed in some way to account for how individuals make decisions, as well as for the practices of market actors. The difficulty of this approach lies in determining how the requirements for consent should be adjusted to empower data subjects. Schremer et al., for instance, have suggested lowering the general legal standard for consent, and reserving the higher standard for situations which 'involve serious risks or consequences for the person who consents'.[125] The insurmountable hurdle of this approach would be how to determine what involves serious risk. Others have suggested developing technical means of providing and revoking consent, similar browser settings for cookies, or using other technical means to help users manage the consent fatigue experienced online.[126] However, it is questionable whether such technical solutions would be sufficient to address the underlying issue, namely that individuals are constantly asked for consent in the online environment while we know that individual decision-making is far from fully-informed or rational.

The two approaches discussed so far have been presented as separate but, in reality, they can be combined and it is possible to propose ways forward for consent. From the analysis in section 3, it seems clear that the requirement for informed consent can be made more specific and precise, to ensure that data subjects actually understand how

124 *Data protection: Commission adopts new rules to ensure stronger enforcement of the GDPR in cross-border cases*, EU Commission (4 July 2023), <https://ec.europa.eu/commission/presscorner/detail/en/ip_23_3609>.

125 Schermer, B. W., Custers, B., & Van der Hof, S. (2014). The crisis of consent: How stronger legal protection may lead to weaker consent in data protection. *Ethics and Information Technology*, *16*(2), 171-182.

126 Custers, B., Fosch-Villaronga, E., van der Hof, S., Schermer, B. W., Sears, A. M., & Tamò-Larrieux, A. (2022). The Role of Consent in an Algorithmic Society–Its Evolution, Scope, Failings and Re-conceptualization. In: Kosta, E., Leenes, R., & Kamara, I. *(2022) Research Handbook on EU data protection. Edward Elgar Publishing*, 455-473.

their data are processed. This can be complemented with further guidance from the EDPB on transparency obligations, especially on the substance of information to be provided as per Articles 13 and 14 GDPR, since informed consent relies on the information provided by the controller. However, such changes should not be made at the expense of addressing the practical issues surrounding enforcement. These among others require that we have a thorough understanding of how the practices of market actors undermine the transparency and consent requirements. Accordingly, it seems that tinkering with the requirements for consent together with bolstering enforcement might bring consent closer to being able to fulfil its empowering role in the online context. A combination of amendments to consent in EU data protection law and a focus on market actors' practices and enforcement might leave hope that consent can still be an empowerment tool. Yet what exactly needs to be amended and how enforcement issues can be remedied continues to be an enigma. The former entails issues of how to amend existing law to ensure specific outcomes, as well as the upholding of values like legal certainty, while avoiding unintended consequences. The latter entails issues of a practical, budgetary and political nature. This chapter has sought to mainly focus on the practices of market actors and show how they undermine the legal requirements, although some suggestions have been made here on how the standard of informed consent could be strengthened in light of these practices.

A third and final approach to how consent can fulfil its empowering role necessarily takes the debate outside of EU data protection law altogether and questions the very business model of very large platforms and much of the internet, which seems geared towards collecting as much personal data as possible for economic exploitation. This approach would see consent continue to struggle to fulfil its empowering role providing the business model of market actors on the internet is not fundamentally challenged. This is because the business model is arguably incompatible with the rights to privacy and data protection.[127] This approach would also require questioning whether it is possible or even desirable to seek to empower individuals to take control over the uses of their personal data in the online context, considering structural issues like power asymmetries between individuals and very large platforms. Perhaps in light of an empirical understanding of market actors' practices, the discussion can be framed in terms of protecting data subjects rather than empowering them. An obvious way in which this might translate into EU data protection law would be to not allow very large platforms to rely on consent in specific circumstances, or to generally move away from consent, which has been already suggested. The danger of framing the discussion in this way would be that the 'hot potato' question of the business model would be moved

127 *See, e.g.* Custers, B., & Malgieri, G. (2022). Priceless data: Why the EU fundamental right to data protection is at odds with trade in personal data. *Computer Law & Security Review*, *45*, 105683.

to other legal bases. This would risk leaving the question of the desirability of the business model unaddressed. New legislative instruments such as the Data Services Act and Digital Markets Act aim to inter alia tackle some of the data practices of very large platforms, but where they defer to the GDPR notion of consent, it remains doubtful that they can empower individuals and challenge the business model. As such, it seems that one cannot conclude a reflection on the role of consent in the online context without admitting that solutions outside EU data protection law are necessary. These solutions would be geared towards challenging the business model of data accumulation and its economic exploitation, not necessarily or exclusively based on the fundamental right to data protection, but also on the basis of how ethical and just the EU citizenry deems such practices.

CHAPTER

# VIII

# The European Data Act: A Horizontal Building Block for the Data Economy?

**Thomas Tombal &**

**Inge Graef**[1]

https://doi.org/10.26116/w7xa-8288

1 At the time of drafting, Thomas Tombal was a postdoctoral researcher at the Tilburg Institute for Law, Technology and Society (TILT) and the Tilburg Law and Economics Center (TILEC) of Tilburg University, and lecturer at the Université de Namur. At the time of publication, Thomas Tombal works as a Case Handler at the European Commission – DG Competition. The views and opinions expressed herein are personal and do not necessarily reflect those of the European Commission or other EU institutions. Inge Graef is Associate Professor Tilburg Institute of Law, Technology, and the Society (TILT), Tilburg Law School (TLS), Tilburg University.

## 1. Introduction

The research project underlying this book seeks to study the shift from human-centric regulation to data-centric regulation. The adoption of the Data Governance Act[2] and the Data Act[3] are the most prominent illustrations of how this shift has accelerated in the European context. The Data Act is of particular relevance considering its broad and ambitious scope. Indicative, in this regard, is that the Data Act was heralded as the 'last horizontal building block of the Commission's data strategy'.[4] At the same time, it is a piece of legislation that adds up to an array of existing legislations that aim to regulate

2 Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), *OJ L 152/1*, 3 June 2022.

3 Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act), *OJ L 1/71*, 22 December 2023. For detailed comments of the Data Act, see Kerber, W. (2023). Governance of IoT data: why the EU Data Act will not fulfill its objectives. *GRUR International*, *72*(2), 120-135; Drexl, J., Banda, C., Gonzalez Otero, B., Hoffmann, J., Kim, D., Kulhari, S.,... & Wiedemann, K. (2022). Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act). <https://www.ip.mpg.de/en/research/research-news/position-statement-on-the-eu-data-act.html>; Picht, P. G. (2023). Caught in the acts: framing mandatory data access transactions under the data act, further EU digital regulation acts, and competition law. *Journal of European Competition Law & Practice*, *14*(2), 67-82, available at <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4076842>, p. 19-42; Colangelo, G. (2022). European Proposal for a Data Act–A First Assessment. *CERRE Evaluation Paper*, available at <https://cerre.eu/wp-content/uploads/2022/07/200722_CERRE_Assessment-Paper_DataAct.pdf>; Habich, E. (2022). FRAND Access to Data: Perspectives from the FRAND Licensing of Standard-Essential Patents for the Data Act Proposal and the Digital Markets Act. *IIC-International Review of Intellectual Property and Competition Law*, *53*(9), 1343-1373; Podszun, R., & Offergeld, P. (2022). The EU data act and the access to secondary markets, *Study for the Ludwig-Fröhler-Institut für Handwerkswissenschaften*, 2022, <https://ssrn.com/abstract=4256882>; Metzger, A., & Schweitzer, H. (2022). Shaping markets: A critical evaluation of the draft data act. <https://ssrn.com/abstract=4222376>; Martens, B. (2023). Pro- and anti-competitive provisions in the proposed European Union Data Act, *Bruegel Working Paper 01/2023*, <https://www.bruegel.org/sites/default/files/2023-01/WP%2001.pdf>.

4 European Commission, 'Data Act: Commission proposes measures for a fair and innovative data economy' (press release), available at <https://ec.europa.eu/commission/presscorner/detail/en/ip_22_1113>. See also Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, '*A European strategy for data*', 19 February 2020, COM(2020) 66.

the European data economy, whether more horizontal,[5] sectoral[6] or targeted at specific digital actors.[7] Against this background, the chapter critically reflects on the Data Act and its ability to regulate the European data economy effectively. While our analysis goes into the details of the different provisions, our purpose is to show whether the design of key aspects of the Data Act allow it to reach its overall objective of making more data available for use in the EU. As such, this chapter builds up to a wider reflection on the effectiveness of the Data Act for stimulating the European data economy.

We argue that, although the Data Act clarifies important conditions applicable to data sharing more generally and can thereby stimulate the European data economy through the additional clarity provided, its scope is more limited than it may appear at first. The Data Act also leaves certain key issues regarding enforcement and its interaction with other laws unaddressed. On the one hand, this may be a choice of the legislator to channel ambitions considering that the Data Act is a first attempt at a horizontal and legally-binding legislative instrument to regulate the sharing of personal as well as non-personal data. On the other hand, this may be due to the fluid nature of data, the

5 See 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 (General Data Protection Regulation), *OJ L 119*, 4 May 2016; Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union, *OJ L 303/59*, 28 November 2018; Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the reuse of public sector information, *OJ L 172/56*, 26 June 2019; Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), *OJ L 152/1*, 3 June 2022.

6 See, for instance, Directive (EU) 2015/2366 of the European Parliament and of the Council of 25 November 2015 on payment services in the internal market, amending Directives 2002/65/EC, 2009/110/EC and 2013/36/EU and Regulation (EU) No 1093/2010, and repealing Directive 2007/64/EC, *OJ L 337/35*, 23 December 2015; Regulation (EU) 2018/858 of the European Parliament and of the Council of 30 May 2018 on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles, amending Regulations (EC) No 715/2007 and (EC) No 595/2009 and repealing Directive 2007/46/EC, *OJ L 151/1*, 14 June 2018; Directive (EU) 2019/944 of the European Parliament and of the Council of 5 June 2019 on common rules for the internal market for electricity and amending Directive 2012/27/EU, *OJ L 158/125*, 14 June 2019.

7 See Directive 2019/770 of the European Parliament and of the Council of 20 May 2019 on certain aspects concerning contracts for the supply of digital content and digital services, *OJ L 136/1*, 22 May 2019; Regulation (EU) 2019/1150 of the European Parliament and of the Council of 20 June 2019 on promoting fairness and transparency for business users of online intermediation services, *OJ L 186/57*, 11 July 2019; Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector (Digital Markets Act), *OJ L 265/1*, 14 September 2022; Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act), *OJ L 277/1*, 19 October 2022.

characteristics and requirements of which change according to the context. In addition, due to its fluid nature, the interactions of data with laws in other areas cannot be fully predicted and clarified upfront beyond a case-by-case setting. As such, the Data Act shows both the potential of regulating a subject matter like data in a horizontal manner as well as its limits, in particular in the form of the remaining need to interpret open questions and to act on a sector-specific basis in parallel. While the Commission may portray the Data Act as 'the last horizontal building block', we believe that its adoption will only just be the end of the beginning of the European data economy's regulation.

According to the Commission, the aim of the Data Act is to address issues that slow down the development of the European data economy, such as the insufficient availability of data for reuse, by aiming to create a legal instrument that would enable wider data use across the economy.[8] This would:

> ensure fairness in the digital environment, stimulate a competitive data market, open opportunities for data-driven innovation and make data more accessible for all [and would] lead to new, innovative services and more competitive prices for aftermarket services and repairs of connected objects.[9]

In order to achieve these objectives, the Data Act combines two approaches. On the one hand, the Data Act imposes specific new data access obligations on businesses, including in the context of Internet of Things (IoT) devices,[10] regulates the sharing of data to public sector bodies in cases of exceptional need[11] and facilitates the switching of data processing services in the cloud environment.[12] On the other hand, the Data Act creates a baseline horizontal framework for compulsory data sharing. This chapter focuses on the latter.

The baseline horizontal framework for compulsory data sharing is at the heart of section 2 of this chapter in particular. It concerns the basic common rules that will need to be applied in any situation where a legislation adopted or revised after the Data Act makes it compulsory for a data holder to make data available to a data recipient.[13] Accord-

8 Commission Staff Working Document, 'Impact assessment report accompanying the document Proposal for a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act)', Brussels, 23 February 2022, SWD(2022) 34 final, available at <https://eur-lex.europa.eu/legal-content/FR/TXT/?uri=SWD%3A2022%3A34%3AFIN&qid=1645811711252>, p. 1 and 7.

9 European Commission, 'Data Act: Commission proposes measures for a fair and innovative data economy', *op. cit.*

10 Articles 4 and 5 of the Data Act.

11 Articles 14 to 22 of the Data Act.

12 Articles 23 to 31 of the Data Act.

13 Article 12.1 of the Data Act. See also Articles 8 to 13. "Data recipient' means a legal or natural person, acting for purposes which are related to that person's trade, business, craft or profession, other than the user of a product or related service, to whom the data holder makes data available, includ-

ing to its Explanatory Memorandum, the key objective underlying the Data Act's baseline horizontal framework is to contribute to:

> the creation of a cross-sectoral governance framework for data access and use by legislating on matters that affect relations between data economy actors, in order to provide incentives for horizontal data sharing across sectors.[14]

As this idea of having common horizontal ground rules is a truly novel approach compared to the previous disparate and mostly sector-specific data sharing approaches, we will analyse how the EU legislator has chosen to build this framework, how it aligns with other existing legislation and whether there are still potential gaps.

Section 3 focuses on enforcement, which is relevant for the Data Act's horizontal framework for compulsory data sharing and for the specific new data access obligations it imposes on businesses. The issue of enforcement is only briefly addressed in the Data Act.[15] However, we believe that many uncertainties result from the text, which could hamper the Data Act's objective of fostering more data sharing in the EU. The provisions could lead to scattered options in terms of the designation of authorities to enforce the Data Act.[16] This could complicate alignment and cooperation between Member State authorities and between authorities within the same Member States, as more than one authority can be designated to enforce this Act.[17] Since the Member State of the entity's main establishment is responsible for monitoring compliance with the Data Act in cross-border cases,[18] strategic behaviour may occur that in the worst case could lead to a similar enforcement bottleneck as in the GDPR.[19]

Section 4 concludes by summarising our main findings pertaining to the two key aspects of the Data Act. It critically reflects on whether its provisions are suitable to achieve the overall objectives it pursues. In this regard, we also provide an outlook to the future of regulating data sharing in the EU.

ing a third party following a request by the user to the data holder or in accordance with a legal obligation under Union law or national legislation adopted in accordance with Union law' (Article 2(14) of the Data Act).

14 Data Act Proposal, Explanatory Memorandum, p. 1.

15 See Articles 37 to 42 of the Data Act.

16 See Article 37 of the Data Act.

17 See Article 37.1 of the Data Act.

18 Article 37.10 of the Data Act.

19 Communication from the Commission to the European Parliament and the Council, '*Data protection as a pillar of citizens' empowerment and the EU's approach to the digital transition – two years of application of the General Data Protection Regulation*', 24 June 2020, COM(2020) 264 final, p. 6; Gentile, G., & Lynskey, O. (2022). Deficient by design? The transnational enforcement of the GDPR. *International & Comparative Law Quarterly*, *71*(4), 799-830.

## 2. The Data Act's Baseline Horizontal Framework for Compulsory Data Sharing

The key objective of the Data Act – to create a cross-sectoral (horizontal) framework for compulsory data sharing – might seem quite broad at first. The Data Act can potentially be understood as applying to any type of compulsory data sharing between businesses and their users (B2U data sharing), between businesses (B2B data sharing), between businesses and public sector bodies (Business-to-Government (B2G) and Government-to-Business (G2B) data sharing), and between public bodies (Government-to-Government (G2G) data sharing). However, the scope of this horizontal framework is much more limited, as we outline below. Beyond this, it is clear that legislation that organises voluntary data sharing, such as the Data Governance Act,[20] will remain unaffected by the Data Act's horizontal framework.[21]

### 2.1 Scope of the Framework

The Data Act's horizontal framework only applies where a data holder is obliged to make data available to a data recipient.[22] Data holders are the actors who have de facto or de jure control over data generated by products or services, and who have the obligation under EU law to make available certain data.[23] This suggests that only businesses that collect data through their control on products or services fall within the scope of this horizontal framework. Accordingly, this would exclude G2B and G2G data sharing from this framework. Here, we can see a clear distinction with the Data Governance Act, which contains a broader definition of 'data holders' as public sector bodies are also covered.[24]

---

20 Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act), *OJ L 152/1*, 3 June 2022.

21 Recital 42 of the Data Act.

22 Article 12.1 of the Data Act.

23 Article 2(13) and Recital 5 of the Data Act.

24 See Article 2(8) of the Data Governance Act: "data holder' means a legal person, including public sector bodies and international organisations, or a natural person who is not a data subject with respect to the specific data in question, which, in accordance with applicable Union or national law, has the right to grant access to or to share certain personal data or non-personal data'. According to the European Data Protection Board (EDPB) and the European Data Protection Supervisor (EDPS), this discrepancy between the two regulations could create some confusion and legal uncertainty. EDPB-EDPS, *Joint Opinion 2/2022 on the Proposal of the Proposal of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act)*, 4 May 2022, available at <https://edpb.europa.eu/our-work-tools/our-documents/edpbedps-joint-opinion/edpb-edps-joint-opinion-22022-proposal-european_en>, p. 11.

Data recipients are defined by the Data Act as 'legal or natural person[s], acting for purposes which are related to that person's trade, business, craft or profession, other than the user of a product or related service'.[25] This would thus exclude B2G data sharing (as public sector bodies do not act for purposes related to trade, business, craft or profession) and B2U data sharing (as users are explicitly excluded) from the scope of this horizontal framework. It therefore only applies to compulsory B2B data sharing, which is confirmed by the Commission itself.[26] Note, however, that B2B data sharing can take place at the initiative of the user when invoking a right to move data from one business to another business.

Looking at the three data sharing obligations contained in the Data Act as mentioned in the introduction, this means the horizontal framework only applies to the IoT data access right to the extent that it benefits third parties,[27] but not to the extent that it benefits the users of connected products or related services,[28] as that is a B2U data sharing scenario. Furthermore, it does not apply to data sharing for exceptional need,[29] as this constitutes B2G data sharing. It does not apply to data portability obligations imposed on providers of data processing services either,[30] as the Data Act does not provide for the possibility of direct data sharing between the provider and a third party. Instead, it only provides for data sharing between the provider and its users (B2U data sharing). These cases fall outside of the horizontal framework and thus require targeted rules.

Importantly, the horizontal framework also does not apply to all legislation imposing B2B data sharing. Instead, it only applies to obligations to make data available under EU legislation or national legislation that implements EU law that will enter into force after 11 January 2024.[31] Therefore, this framework will not apply to any pre-existing (sectoral) legislation imposing compulsory B2B data sharing, such as Article 20.2 GDPR.

The Data Act's horizontal framework for compulsory data sharing is thus much narrower than it might appear at first sight. Accordingly, some uncertainty might remain

25 Article 2(14) of the Data Act. According to the EDPB and the EDPS, the Data Act should specify whether data intermediation services, as defined in Art. 2(11) of the Data Governance Act, are covered as recipients, as Recital 35 of the Data Act Proposal seems to suggest (EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal*, *op. cit.*, p. 11-12).

26 Commission Staff Working Document, 'Impact assessment report accompanying the Data Act', *op. cit.*, p. 68.

27 Article 5 of the Data Act.

28 Article 4 of the Data Act.

29 See Articles 14 to 22 of the Data Act.

30 See Articles 23 to 31 of the Data Act.

31 Article 12.1 and 44.1 of the Data Act.

in the coming years due to the parallel application of different rules and regimes. This is not tenable in the long term, and some convergence between the Data Act's horizontal framework and other data sharing instruments will have to be achieved in the future.

Finally, it is important to outline that this is a baseline horizontal framework. This implies that the rules it contains lay down the minimum standards applicable to any legislations that fall within its scope. However, it does not prevent the adoption of more specific and far-reaching rules in the context of individual sectors or that of the development of common European data spaces.[32] These more specific rules could, for example, include further-reaching requirements on technical aspects of the data access,[33] on limits to the data holder's right to use certain data provided by users or on other aspects that go beyond data access matters.[34] In this regard, it will be important to find the right balance between accommodating sector-specific needs and maintaining a minimum level of coherence between the different instruments to avoid the creation of a patchwork of data sharing regimes that would be hard to navigate.

## 2.2 Framework Design

### 2.2.1 The Central Role of Data Sharing Agreements

The cornerstone of this horizontal framework is that data holders have to share the data with data recipients in a transparent manner and under fair, reasonable and non-discriminatory (FRAND) terms, through the means of a data sharing agreement.[35] Importantly, the Data Act underlines that the principle of contractual freedom must be respected. Accordingly, it will be up to the parties involved to negotiate these terms and to agree on what constitutes FRAND terms in their particular situation.[36] Some argue that this contractual path is questionable as 'it increases transaction costs for data recipients who cannot obtain the data directly via an open interface or in other auto-

32 Article 44.2 and Recital 115 of the Data Act. See, for instance: Proposal for a Regulation of the European Parliament and of the Council on the European Health Data Space, 3 June 2022, COM(2022) 197 final.

33 Such as 'interfaces for data access, or how data access could be provided, for example directly from the product or via data intermediation services' (Recital 115 of the Data Act).

34 Article 44.2 of the Data Act.

35 Articles 8.1 and 8.2 of the Data Act. According to the EDPB and the EDPS, the fact that the data subject, whose data might be shared, plays no role in the elaboration of such contract 'risks to severely compromise the effectiveness of data protection rights' (EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal*, *op. cit.*, p. 17). For some, this unilateral formulation against the data holder is problematic and the data recipients should also be under the obligation to negotiate FRAND terms in good faith (see Drexl, J., Banda, C., Gonzalez Otero, B., Hoffmann, J., Kim, D., Kulhari, S.,... & Wiedemann, K. (2022). Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act) *op. cit.*, p. 39).

36 Recital 43 of the Data Act.

mated ways'.[37] This is because, while 'negotiations may lead to efficient results in the B2B area if the companies are on the same level and there are no power asymmetries', requiring a contractual agreement creates 'considerable potential for disrupting free access to data, since remuneration and conditions are to be negotiated by parties with very different bargaining positions'.[38]

Despite this, the parties' contractual freedom is limited to the extent that the data sharing agreement terms may not exclude the application of this horizontal regime, derogate from it or vary its effect.[39] These rules are imperative, and any contractual term that does not comply with them shall not be binding on the party to which it is detrimental.[40]

### 2.2.2 FRAND Requirements

#### 2.2.2.1 *Non-Discriminatory*

In terms of the FRAND requirements, the Data Act provides that the data holder shall not discriminate (i.e. share data at more favourable conditions) between comparable categories of data recipients, including partner and linked enterprises[41] (i.e. engage in

37 Podszun, R., & Offergeld, P. (2022). The EU data act and the access to secondary markets. *Available at SSRN 4256882*, *op. cit*., p. 27; Drexl, J., Banda, C., Gonzalez Otero, B., Hoffmann, J., Kim, D., Kulhari, S.,... & Wiedemann, K. (2022). Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act), *op. cit*., p. 29.

38 Podszun, R., & Offergeld, P. (2022). The EU data act and the access to secondary markets. *Available at SSRN 4256882*, *op. cit*., p. 28. See p. 28-29 for further discussions on this need for a more direct access. See also J. Drexl, C. Banda, B. González Otero, J. Hoffmann, D. Kim, S. Kulhari, V. Moscon, H. Richter and K. Wiedemann, *Position Statement of the Max Planck Institute for Innovation and Competition on the Commission's Data Act Proposal*, *op. cit*., p. 28.

39 Article 8.2 of the Data Act.

40 Ibid.

41 "Linked enterprises' are enterprises which have any of the following relationships with each other: (a) an enterprise has a majority of the shareholders' or members' voting rights in another enterprise; (b) an enterprise has the right to appoint or remove a majority of the members of the administrative, management or supervisory body of another enterprise; (c) an enterprise has the right to exercise a dominant influence over another enterprise pursuant to a contract entered into with that enterprise or to a provision in its memorandum or articles of association; (d) an enterprise, which is a shareholder in or member of another enterprise, controls alone, pursuant to an agreement with other shareholders in or members of that enterprise, a majority of shareholders' or members' voting rights in that enterprise' (Article 3.3 of the Annex to the Commission Recommendation 2003/361/EC of 6 May 2003 concerning the definition of micro, small and medium-sized enterprises, *OJ L 124/36*, 20 May 2003); "Partner enterprises' are all enterprises which are not classified as linked enterprises (...) and between which there is the following relationship: an enterprise (upstream enterprise) holds, either solely or jointly with one or more linked enterprises (...) 25 % or more of the capital or voting rights of another enterprise (downstream enterprise)' (Article 3.2 of the Annex to Recommendation 2003/361/EC).

self-preferencing[42]).[43] Importantly, it is up to the data holder to demonstrate that a contractual term is not discriminatory, by proving that a potential difference between two contracts with similar recipients is justified by objective reasons.[44] In this regard, differentiating between recipients from distinct sectors could be justified, as this would 'preserve market-driven incentives to invest in secondary markets with below-average potential of value creation, warranting preferential access conditions for the same input data as other secondary markets'.[45] Some suggest that, in order to verify whether discrimination has occurred, 'it could be required that all previous data access contracts of a company be disclosed to [a specific body]', and that these previous contracts 'could be inspected upon request by arbitration bodies or courts'.[46] Finally, a data holder cannot share data with a data recipient on an exclusive basis unless this has been explicitly requested by a user of the data holder's product or service.[47]

#### 2.2.2.2 *Reasonable*

The Data Act provides that any compensation agreed between the data holder and the data recipient should be reasonable,[48] unless the specific legislation imposing data sharing excludes any compensation at all or provides for a lower one in justified cases.[49] The underlying idea is to incentivise the data holder's 'continued investment

42 On prohibitions of self-preferencing, see also GCEU, *Google Shopping,* 10 November 2021, T-612/17, EU:T:2021:763; Digital Markets Act, *op. cit*., Article 6.5.

43 Article 8.3 of the Data Act.

44 Article 8.3 and Recital 45 of the Data Act.

45 Habich, E. (2022). FRAND Access to Data: Perspectives from the FRAND Licensing of Standard-Essential Patents for the Data Act Proposal and the Digital Markets Act. *IIC-International Review of Intellectual Property and Competition Law*, *53*(9), 1343-1373, p. 1354.

46 Podszun, R., & Offergeld, P. (2022). The EU data act and the access to secondary markets, p. 36.

47 Article 8.4 of the Data Act. As outlined by the EDPB and the EDPS, these non-discrimination and non-exclusivity obligations should however not 'undermine the right of informational self-determination of data subjects according to which they are entitled to discriminate among the recipients of their personal data' (EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal, op. cit*., p. 17).

48 Article 9.1, 9.2 and 9.3 of the Data Act. For a more detailed analysis of what a 'reasonable compensation' could entail, see Monti, G., Tombal, T., & Graef, I. (2022). Study for developing criteria for assessing "reasonable compensation" in the case of statutory data access right: Study for the European Commission Directorate-General Justice and Consumers. <https://data.europa.eu/doi/10.2838/19186>; Habich, E. (2022). FRAND Access to Data: Perspectives from the FRAND Licensing of Standard-Essential Patents for the Data Act Proposal and the Digital Markets Act. *IIC-International Review of Intellectual Property and Competition Law*, *53*(9), 1343-1373, p. 1354-1371. It is also worth noting that the option not to include any compensation for the data holder has been explored but was eventually not retained, due to fears that this would unduly affect the data holders' business interests (Commission Staff Working Document, 'Impact assessment report accompanying the Data Act', *op. cit*., p. 36).

49 Article 9.6 and Recital 50 of the Data Act. Justified cases include 'including the need to safeguard consumer participation and competition or to promote innovation in certain markets' (Recital 50). For instance, the Digital Markets Act provides for a free data portability right (Commission

in generating and making available valuable data, including investments in relevant technical tools'.[50] In fact, the Data Act itself provides for a lower compensation in situations where the data recipient is a micro, small or medium enterprise (SME).[51] The goal is to protect them from excessive economic burdens that would hamper the development of innovative business models.[52] Accordingly, in B2b (the lower case 'b' referring to SMEs) data sharing scenarios, the compensation should not exceed the costs directly related to making the data available[53] and which are attributable to the request.[54] Importantly, this should not be understood as paying for the data itself,[55] but only for the costs incurred for making the data available 'including, in particular, the costs necessary for the formatting of data, dissemination via electronic means and storage'.[56] The costs may vary depending on the arrangements taken for making the data available:

> Long-term arrangements between data holders and data recipients, for instance via a subscription model or the use of smart contracts, may reduce the costs in regular or repetitive transactions in a business relationship. Costs related to making data available are either specific to a particular request or shared with other requests. In the latter case, a single data recipient should not pay the full costs of making the data available.[57]

The 'reasonable compensation' to be paid by a recipient that is not an SME will necessarily include these direct sharing costs plus a fee covering (at least in part) 'investments in the collection and production of data, where applicable, taking into account whether

Staff Working Document, 'Impact assessment report accompanying the Data Act', *op. cit*., p. 127). Such a distinction can arguably be justified by the fact that the DMA only applies to a limited set of very powerful 'gatekeepers', whose interests should arguably weight less heavily in the balance in light of the market failures they strive on and of the need to foster contestability on those markets.

50 Recital 46 of the Data Act.

51 As defined in Article 2 of the Annex to Recommendation 2003/361/EC.

52 Recital 49 of the Data Act.

53 "Directly related costs are those costs which are attributable to the individual requests, taking into account that the necessary technical interfaces or related software and connectivity will is to be established on a permanent basis by the data holder' (Recital 49 of the Data Act).

54 Article 9.4 and 9.2.a) of the Data Act.

55 Recital 46 of the Data Act. The EDPB and the EDPS have expressed concerns about this reference to 'paying for the data itself', as it would acknowledge the possibility to monetise personal data, while 'data protection is a fundamental right guaranteed by Article 8 of the Charter and personal data cannot be considered as a tradeable commodity' (EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal*, *op. cit*., p. 18).

56 Article 9.2.a) of the Data Act.

57 Recital 47 of the Data Act.

other parties contributed to the obtaining, generating or collecting [of] the data in question'.[58] In addition, the compensation 'may also depend on the volume, format and nature of the data'.[59] The reasonable nature of the compensation will thus be a function of the prevailing market conditions and may include a margin, except for SMEs and not-for-profit research organisations.[60] This latter aspect is in line with the Open Data Directive,[61] which further defines a 'reasonable return on investment' as 'a percentage of the overall charge, in addition to that needed to recover the eligible costs, not exceeding 5 percentage points above the fixed interest rate of the [European Central Bank]'.[62] The Data Act is not as specific, which is logical considering the much larger range of situations and types of data it covers in comparison with the Open Data Directive, but it does contain some useful guidance in the recitals. The margin 'may consider the costs for collecting the data'. It may also:

> decrease where the data holder has collected the data for its own business without significant investments or may increase where the investments in the data collection for the purposes of the data holder's business are high. It may be limited or even excluded in situations where the use of the data by the data recipient does not affect the data holder's own activities. The fact that the data is co-generated by a connected product owned, leased or rented by the user could also lower the amount of the compensation in comparison to other situations where the data are generated by the data holder for example during the provision of a related service.[63]

In any case, the data holder will have to provide sufficiently detailed information to the recipient setting out the basis for the calculation of the requested compensation so that the latter can verify that the required compensation complies with the Data Act.[64] This suggests that the initial offer should come from the data holder, allowing the data recipient to make a counter-offer. In this regard, the negotiation framework for the licensing of Standard-Essential-Patents (SEPs) on FRAND terms, as proposed in the Huawei case,[65] could be used as an inspiration to assist the parties in reaching an agree-

58 Article 9.2.b) of the Data Act.
59 Article 9.3 of the Data Act.
60 Recital 47 of the Data Act. See also Commission Staff Working Document, 'Impact assessment report accompanying the Data Act', *op. cit.*, p. 154.
61 Article 6.4 of Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the reuse of public sector information, *OJ L 172/56*, 26 June 2019.
62 Article 2(16) of Directive (EU) 2019/1024.
63 Recital 47 of the Data Act.
64 Article 9.7 and Recital 51 of the Data Act.
65 ECJ, *Huawei*, 16 July 2015, C-170/13, EU:C:2015:477, §§ 60-69.

ment on the remuneration.[66] If no agreement is reached following the counter-offer, the parties should, by common agreement, request that the price be determined by an independent third party (which could be a designated entity [e.g. the Support Centre for Data Sharing[67]]) or a dispute settlement body (e.g. supervising authorities, arbitrators or courts).

An important question underlying this negotiation between the data holder and the recipient is that of the relationship between 'the data holder's right to demand a FRAND compensation for data access [under Art. 8.1] and its access duty to a third party at the request of the user without undue delay [under Art. 5.1]'.[68]

> Allowing the data holder to retain the data until the FRAND dispute is resolved would lead to a violation of the obligation of the data holder vis-à-vis the user and seriously affect the effectiveness of the data access and use right of the latter. Conversely, if one considers the data holder under an obligation to provide access despite its failure to agree on FRAND terms, this would create a so-called 'hold-out' situation, where the third party can simply refuse or evade honest FRAND negotiations, as this will not hinder the provision of the service.[69]

Surprisingly, the Data Act is silent on this tension, even though it is fundamental to ensure that both the holder and the recipient engage in 'diligent and good faith dealing'.[70]

As such, we argue that a balance needs to be found between the interests of data holders and data recipients to ensure effective negotiations. This can be done in different ways. Some have suggested to provide the recipient with immediate access to the

66 Wendehorst, C., & Cohen, N. (2023). ALI/ELI Principles for a Data Economy: Data Transactions and Data Rights. <https://europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ALI-ELI_Principles_for_a_Data_Economy_Final_Council_Draft.pdf>, p. 177. On this point, see also Drexl, J. (2017). Designing competitive markets for industrial data. *J. Intell. Prop. Info. Tech. & Elec. Com. L.*, *8*, 257, p. 55; Habich, E. (2022) 'FRAND Access to Data', *op. cit.*, p. 1361-1370; Picht, P. 'Caught in the acts', *op. cit.*, p. 34-36; Tombal, T. (2020). Economic dependence and data access. *IIC-International Review of Intellectual Property and Competition Law*, *51*(1), 70-98, p. 94; Tombal, T. (2022). *Imposing Data Sharing among Private Actors: A Tale of Evolving Balances*, Alphen aan den Rijn, Kluwer Law International (Information Law Series, Vol. 48), p. 298-299.

67 See <https://eudatasharing.eu/>.

68 Habich, E. (2022) 'FRAND Access to Data', *op. cit.*, p. 1358.

69 Drexl, J., Banda, C., Gonzalez Otero, B., Hoffmann, J., Kim, D., Kulhari, S.,... & Wiedemann, K. (2022). Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act), p. 28.

70 Picht, P. G. (2023). Caught in the acts: framing mandatory data access transactions under the data act, further EU digital regulation acts, and competition law. *Journal of European Competition Law & Practice*, *14*(2), 67-82, *op. cit.*, p. 35.

data in exchange for the deposit of a lump sum.[71] One alternative is to take away the recipient's access to the data if it is found to engage in a manifestly uncooperative behaviour in order to avoid 'hold out' tactics.[72] At the same time, the data holder should not be able to raise a lack of agreement on the 'reasonable compensation' as a way to bypass its duty to share. This would incentivise the data holders to drag their feet and to engage in bad faith negotiations with the recipient by abusing delaying tactics. Even though the Data Act does not specify any such measures, there is still scope to implement the relevant provisions in a way that balances the interests of data holders and data recipients.

In addition, it cannot be excluded that any remuneration charged to data recipients will indirectly affect users as the price paid by third parties to access the data will likely be passed on to them in the cost of the product or service.[73] As this runs counter to the principle according to which data sharing between a holder and a third party should not entail any financial burdens for the user,[74] some argue that the data holder should not be remunerated by third parties either.[75] This would also have the merit to supress the above-mentioned 'negotiation hold-out' issue. The recitals to the Data Act do explicitly leave room for excluding any margin on the part of the data holder (beyond the costs for making the data available under Article 9.2.a) of the Data Act) 'in situations where the use of the data by the data recipient does not affect the own activities of the data holder'.[76] It remains to be seen whether data recipients indeed can convince data holders not to charge any extra margins. This issue is worth monitoring closely, so it is beneficial that the final text of the Data Act requires the Commission to adopt guidelines on the calculation of reasonable compensation in which such issues can be further elaborated.[77]

71 Podszun, R., & Offergeld, P. (2022). The EU data act and the access to secondary markets, *op. cit.*, p. 36.

72 See Habich, E. (2022). 'FRAND Access to Data', *op. cit.*, p. 1360-1361 and 1368-1370.

73 Martens, B. (2023). Pro-and anti-competitive provisions in the proposed European Union Data Act, *op. cit.*, p. 12.

74 Article 5.1 of the Data Act.

75 J Drexl, J., Banda, C., Gonzalez Otero, B., Hoffmann, J., Kim, D., Kulhari, S.,... & Wiedemann, K. (2022). Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act), *op. cit.*, p. 38; Martens, B. (2023). Pro-and anti-competitive provisions in the proposed European Union Data Act, *op. cit.*, p. 11; Podszun, R., & Offergeld, P. (2022). The EU data act and the access to secondary markets, *op. cit.*, p. 37-38.

76 Recital 47 of the Data Act.

77 Article 9.5 of the Data Act.

#### 2.2.2.3 *Fair*

Moving on to the fair nature of the terms, it must be underlined that 'fairness' is not specifically defined in the Data Act, although an 'unfairness test' has been created for situations where a 'contractual term, concerning the access to and use of data or the liability and remedies for the breach or the termination of data related obligations … has been unilaterally imposed by an enterprise on another enterprise'.[78] The Data Act outlines several situations in which contractual terms qualify as unfair or are presumed unfair.[79] Beyond this, it is worth outlining that the Data Act provides that the Commission will develop and recommend non-binding model contractual terms that should assist the parties in drafting balanced data sharing agreements.[80] These model terms, which could take into account specific sectoral conditions and existing practices used in the context of voluntary data sharing mechanisms where necessary,[81] should lead to fairer data sharing contracts.[82]

The horizontal framework also provides rules that ensure that the data sharing obligation does not affect the technical and commercial security of the data holder.[83] First, the Data Act provides information about the articulation with the data holder's trade secrets.[84] This shows that this horizontal framework aims to find a fair balance between the holder's business interests and the benefits that the sharing obligation generates for the recipients. The Data Act provides that an obligation to share data with a data recipient should not oblige the disclosure of trade secrets, unless provided otherwise in the legislation imposing the sharing.[85] It should be outlined here that some believe that 'this

78 Article 13 of the Data Act.

79 Article 13.4 and 5 of the Data Act.

80 Article 41 of the Data Act.

81 See the work of the 'Support Centre for Data Sharing' (https://eudatasharing.eu/homepage), which has been created by the Commission in order to put in place a series of measures facilitating voluntary data sharing, in particular by providing examples of good practices, standard contractual clauses or existing contract models (Commission Staff Working Document establishing a guidance on sharing private sector data in the European data economy accompanying the Communication '*Towards a common European data space*', Brussels, 25 April 2018, SWD(2018) 125 final, p. 6).

82 Recital 111 of the Data Act. Metzger and Schweitzer recommend to look for inspiration in the 'ALI/ELI Draft Principles for a Data Economy' (Wendehorst, C., & Cohen, N. (2021). ALI/ELI Principles for a Data Economy: Data Transactions and Data Rights, *ELI Final Council Draft*, <https://europeanlawinstitute.eu/fileadmin/user_upload/p_eli/Publications/ALI-ELI_Principles_for_a_Data_Economy_Final_Council_Draft.pdf>) when drafting these model contract terms (Metzger, A., & Schweitzer, H. (2022). Shaping markets: A critical evaluation of the draft data act. *Available at SSRN 4222376*, p. 19).

83 Article 11 of the Data Act.

84 See Directive (EU) 2016/943 of the European Parliament and of the Council of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure, *OJ L 157*, 15 June 2016.

85 Article 8.6 of the Data Act.

provision opens the door for the data holder to strategically claim the existence of trade secrets to refuse the sharing of the data'.[86] Others outline that Article 8.6 'arguably merely clarifies that the existence of the Data Act does not change the fact that the data is still subject to trade secret protection. It does not restrict the Data Act's data access claims'.[87] The final text of the Data Act further specifies how the balance between trade secret protection and the interest in stimulating data sharing should be struck. In particular, the Data Act now provides that a data holder can withhold or suspend data sharing based on a duly substantiated and written decision when the confidentiality of trade secrets is undermined.[88] Only in exceptional circumstances, the data holder may refuse on a case-by-case basis the request for access to the specific data in question when it can demonstrate that it is highly likely to suffer serious economic damage from the disclosure of trade secrets.[89] In both cases, the data holder has to notify the national competent authority. 'Serious economic damage' implies serious and irreparable economic losses. Based on objective elements, the data holder has to demonstrate 'the concrete risk of serious economic damage expected to result from a specific data disclosure and the reasons why the measures taken to safeguard the requested data are not sufficient'.[90] Such objective elements include 'the enforceability of trade secrets protection in third countries, the nature and level of confidentiality of the data requested, and the uniqueness and novelty of the connected product'.[91] Although it remains to be seen how these stipulations are applied in practice, they provide welcome guidance regarding the factors to be taken into account in assessing claims regarding trade secret protection. It is up to the national competent authorities (under the control of the courts) to develop decision-making practice on this issue. Worth mentioning is the explicit statement in the recitals that '[d]ata holders cannot, in principle, refuse a data access request under this Regulation solely on the basis that certain data is considered to be a trade secret, as this would subvert the intended effects of this Regulation'.[92] However, beyond this outer boundary, the provisions still leave significant room for interpretation.

86 Drexl, J., Banda, C., Gonzalez Otero, B., Hoffmann, J., Kim, D., Kulhari, S.,... & Wiedemann, K. (2022). Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act), *op. cit.*, p. 102.

87 Podszun, R., & Offergeld, P. (2022). The EU data act and the access to secondary markets, *op. cit.*, p. 43.

88 Article 4(7) and 5(10) of the Data Act.

89 Article 4(8) and 5(11) of the Data Act.

90 Recital 31 of the Data Act.

91 Article 4(8) and 5(11) of the Data Act.

92 Recital 31 of the Data Act.

Second, the Data Act provides that the data holder may apply appropriate technical protection measures, including smart contracts,[93] to prevent unauthorised access to the data and to ensure compliance with the Data Act as well as with the agreed contractual terms.[94] Nevertheless, these technical protection measures may not be used to hinder the data sharing obligation.[95] In this regard, some argue that this provision should also 'further specify the technical requirements necessary to enable a legally compliant access and use'.[96] A key issue in this regard is that the Data Act 'does not specify in which format the data must be made accessible',[97] as this might significantly increase transaction costs. Accordingly, a fine line must be found between creating sufficient security and preserving the essence of the data sharing obligation.

Third, the data holder is protected against a data recipient's deceptive or abusive conduct. This covers situations like where the latter has provided inaccurate or false information to the data holder to obtain the data; deployed deceptive or coercive means or has abused evident gaps in the technical infrastructure of the data holder designed to protect the data; used the data for unauthorised purposes; or disclosed the data to another party without the data holder's authorisation.[98] Indeed, in such cases, the recipient must take several measures, including erasure of the data without undue delay, ending the production (or alike) of goods, derivative data or services produced on the basis of knowledge obtained through the data.[99] These seem reasonable requirements to protect data holders' interests.

93 "Smart contract' means a computer program used for the automated execution of an agreement or part thereof, using a sequence of electronic data records and ensuring their integrity and the accuracy of their chronological ordering' (Article 2(39) of the Data Act). It is worth noting that the Data Act imposes five essential requirements (robustness and access control; safe termination and interruption; data archiving and continuity; access control; and consistency with the terms of the data sharing agreement) on vendors of smart contracts and on persons whose trade, business or profession involves the deployment of smart contracts for others in the context of data sharing agreements (Article 36 of the Data Act).

94 Article 11.1 of the Data Act.

95 Ibid.

96 Metzger, A., & Schweitzer, H. (2022). Shaping markets: A critical evaluation of the draft data act, *op. cit*., p. 23. For these authors, '[t]his gap is even more striking in light of the detailed provision on the technical requirements for interoperability in Article 28' (Ibid.)

97 Podszun, R., & Offergeld, P. (2022). The EU data act and the access to secondary markets, *op. cit*., p. 30. These authors suggest 'to include an obligation that data must be provided in a structured, common and machine-readable format or in a manner customary in the sector' (p. 31).

98 Article 11.3 of the Data Act.

99 Ibid.

## 2.3 Articulation with Personal Data Protection Considerations

We believe that an important gap in this horizontal framework is the lack of specific guidance regarding the interaction between data sharing obligations and the need to comply with the GDPR's personal data protection rules in situations where personal data have to be shared.[100] The Data Act states that it is without prejudice to the GDPR and that the GDPR prevails in cases of conflict.[101] One may therefore argue that this already clarifies the relationship between the two interests and that the legislator has set a hierarchy. However, in practice, such an approach is unlikely to be workable as many datasets targeted by the Data Act will include personal data and it will often not be clear at the outset to what extent any possible tension with the GDPR can be reconciled without putting the Data Act aside. Of course, the potential benefits that derive from data sharing are only acceptable if this is done in compliance with the rights of the individuals whose personal data could be shared.[102] However, because of the absence of more specific guidance on this issue in the Data Act itself, data holders might strategically invoke GDPR-compliance considerations to justify refusals to give effect to data sharing requests.[103]

A reason for the silence on this interaction with the GDPR in the Data Act may be that it is easier to strike a balance in each specific legislation imposing data sharing, as the compliance with the purpose limitation[104] and data minimisation[105] principles of the GDPR will be a function of the specific data covered and of the specific circumstances justifying the compulsory sharing. For instance, in the context of the IoT data access right, the Data Act provides that IoT generated personal data can only be shared with a third party if there is a valid legal basis under Article 6 or 9 of the GDPR.[106] In this regard,

100 On this point, see EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal*, *op. cit.*, p. 17-19; Drexl, J., Banda, C., Gonzalez Otero, B., Hoffmann, J., Kim, D., Kulhari, S.,... & Wiedemann, K. (2022). Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act), *op. cit.*, p. 105-111.

101 Article 1(5) of the Data Act.

102 See European Data Protection Supervisor, *Opinion 3/2020 on the European strategy for data*, 16 June 2020, available at <https://edps.europa.eu/sites/edp/files/publication/20-06-16_opinion_data_strategy_en.pdf>, p. 8; Tombal, T. (2022). *Imposing Data Sharing among Private Actors*, *op. cit.*, p. 186-191 and 391-393.

103 Metzger, A., & Schweitzer, H. (2022). Shaping markets: A critical evaluation of the draft data act, *op. cit.*, p. 27.

104 Personal data can only be processed for specified, explicit and legitimate purposes, and cannot be further processed in a manner that is incompatible with those purposes (Article 5.1.b) of the GDPR).

105 Only the adequate, relevant and necessary data for the fulfilment of the specific purpose of processing shall be processed (Article 5.1.c) of the GDPR).

106 Articles 4.12 and 5.7 of the Data Act.

it is worth pointing out that the Data Act underlines that 'this Regulation does not create a legal basis for providing access to personal data or making personal data available to a third party'.[107] Another legal basis, such as consent or legitimate interests of the data controller,[108] thus has to be relied on. Importantly, according to the principle of separate justification, which provides that 'each transaction in data requires a legal basis at two levels: the level of the supplier of the data and the level of the recipient',[109] the third-party recipient will need its own legal basis for the further processing of this personal data.[110]

Regarding the IoT data access right, the Data Act also provides that third parties should only process the data 'for the purposes and under the conditions agreed with the user and subject to Union and national law on the protection of personal data including the rights of the data subject insofar as personal data are concerned', [111] and that they shall not use the data for profiling[112] purposes 'unless it is necessary to provide the service requested by the user'.[113] According to the EDPB and the EDPS, the Data Act should however explicitly remind all that any further personal data processing must comply with Article 6.4 of the GDPR, and should include clearer limitations or restrictions of reuse for 'purposes of direct marketing or advertising, employee monitoring, credit scoring or to determine eligibility to health insurance, to calculate or modify insurance premiums'.[114]

In light of the above, we believe that some overlapping principles are worth developing. An example could be a requirement to share, to the extent possible, anonymised data[115] rather than personal data – notably through resorting to the services of interme-

107 Recital 7 of the Data Act.

108 Respectively Articles 6.1.a) and 6.1.f) of the GDPR.

109 Wendehorst, C. (2017). Of elephants in the room and paper tigers: how to reconcile data protection and the data economy. In *Wendehorst, Of Elephants in the Room and Paper Tigers: How to Reconcile Data Protection and the Data Economy, in Lohsse/Schulze/Staudenmayer (Eds), Trading Data in the Digital Economy: Legal Concepts and Tools: Münster Colloquia on EU Law and the Digital Economy* (Vol. 3, pp. 327-356), p. 334-337.

110 Tombal, T. (2022). *Imposing data sharing among private actors*, *op. cit*., p. 186.

111 Article 6.1 of the Data Act.

112 "Profiling' means 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' (Article 4(4) of the GDPR).

113 Article 6.2.b) of the Data Act.

114 EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal*, *op. cit*., p. 3 and 15-16.

115 The ISO 29100 standard defines anonymisation as the : 'process by which personally identifiable information (PII) is irreversibly altered in such a way that a PII principal can no longer be identified directly or indirectly, either by the PII controller alone or in collaboration with any other party' (ISO 29100:2011, point 2.2, available at <https://www.iso.org/obp/ui/#iso:std:iso-iec:29100:ed-1:v1:en>).

diary 'data trustees' charged with anonymising the data –[116] or a requirement to pseudonymise[117] the data where possible and to the extent that this does not affect reuse possibilities.[118] Moreover, the 'reasonable' compensation to be charged by the data holder to the data recipient could be interpreted as including the marginal costs of anonymising or pseudonymising the data.[119] It should also be recalled that if the recipient's further processing is not compatible with the data holder's initial processing, the recipient can only carry out the desired processing under the initial legal ground for processing if it has obtained the data subjects' consent or if this transfer is necessary to comply with a legal obligation.[120] In this regard, the specific purpose pursued will have to be outlined and only the data that are necessary to achieve this purpose can be shared with the recipient.[121] Finally, the data subjects should be clearly informed about this transfer in order to be able to exercise their rights.[122]

### 2.4 Dispute Resolution Mechanism

This horizontal framework for compulsory B2B data sharing is a welcome development as it aims to clarify upfront how to implement data sharing obligations, while trying to find a balance between the interests of both data holders and data recipients. However, by relying on data sharing agreements and by only limiting the data holder's contractual freedom to the extent necessary, it may inevitably cause delays to the data sharing process. Indeed, one can imagine that the data holder and the data recipient will not necessarily agree on what constitutes FRAND terms in their particular situation, and

116 Podszun, R., & Offergeld, P. (2022). The EU data act and the access to secondary markets, *op. cit*., p. 20.

117 Pseudonymisation is 'the processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organisational measures to ensure that the personal data are not attributed to an identified or identifiable natural person' (Article 4.5 of the GDPR).

118 See, by analogy, Article 5.3 of the Data Governance Act. See also EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal, op. cit*., p. 14-15.

119 See, by analogy, Article 6.1 of Directive 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the reuse of public sector information, *OJ L 172/56*, 26 June 2019.

120 Article 6.4 and Recital 50 of the GDPR. For more details, see De Terwangne, C. (2020). Principles relating to processing of personal data. In *The EU general data protection (GDPR): a commentary* (pp. 309-320). Oxford University Press, p. 309-320; Kotschy, W. (2020). Article 6 Lawfulness of processing. In Kuner, C. Bygrave, L. & Docksey, C. (eds.), *The EU General Data Protection Regulation (GDPR)*. Oxford University Press, Oxford, p. 321-344; European Union Agency for Fundamental Rights and Council of Europe, *Handbook on European data protection law*, Luxembourg, Publications Office of the European Union, 2018, p. 122-125.

121 Articles 5.1.b) and c) of the GDPR.

122 EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal, op. cit*., p. 14-15.

that this could lead to lengthy discussions slowing down the sharing process. This risk is in fact very real, as can be observed from the similar difficulties of determining what constitutes a FRAND licence in the context of standard essential patents (SEPs).[123] Some even argue that the potential for disputes will be higher for data licencing than for SEP licencing.[124]

To alleviate this potential issue, the Data Act establishes a dispute resolution mechanism that should offer 'a simple, fast and low-cost' solution to the parties.[125] More concretely, data holders and data recipients should have access to a dispute settlement body (DSB) that has been certified by the Member State in which it is established.[126] Such certification must be requested by the DSB, which will need to demonstrate that:

> (a) it is impartial and independent, and it is to issue its decisions in accordance with clear, non-discriminatory and fair rules of procedure;
>
> (b) it has the necessary expertise, in particular in relation to fair, reasonable and non-discriminatory terms and conditions, including compensation, and on making data available in a transparent manner, allowing the body to effectively determine those terms and conditions;
>
> (c) it is easily accessible through electronic communication technology;
>
> (d) it is capable of adopting its decisions in a swift, efficient and cost-effective manner in at least one official language of the Union.[127]

123 See, for instance, Geradin, D. (2013). Ten Years of DG Competition Effort to Provide Guidance on the Application of Competition Rules to the Licensing of Standard-Essential Patents: Where Do We Stand?. *Available at SSRN 2204359*., p. 7-8; Graham, C., & Morton, J. (2014). Latest EU Developments in Standards, Patents, and FRAND Licensing. *EURO. INTELL. PROP. REV.*, *36*, 700-705; Stern, R. H. (2015). What Are Reasonable and Non-Discriminatory Terms for Licensing a Standard-Essential Patent?. *Eur. Intell. Prop. Rev.*, *37*, 549-549; Drexl, J. (2017). Designing competitive markets for industrial data. *J. Intell. Prop. Info. Tech. & Elec. Com. L.*, *8*, 257, *op. cit*., p. 55; ECJ, *Huawei*, 16 July 2015, C-170/13, EU:C:2015:477.

124 Drexl, J., Banda, C., Gonzalez Otero, B., Hoffmann, J., Kim, D., Kulhari, S.,... & Wiedemann, K. (2022). Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act), *op. cit*., p. 38.

125 Article 10.1 and Recital 52 of the Data Act. Once again, the EDPB and the EDPS highlight that the data subject, whose data might be shared, is completely overlooked in this dispute resolution mechanism and that this may interfere with her right to lodge a complaint with a supervisory authority. (EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal*, *op. cit*., p. 18).

126 Articles 10.1 and 10.5 of the Data Act.

127 Article 10.5 of the Data Act.

To make it easier for data holders and data recipients to identify these certified DSBs, the Member States will have to notify the Commission about the certified DSBs, so that they can be included in a list published and updated by the Commission on a dedicated website.[128]

With regard to the functioning of these DSBs, it is important to highlight that DSBs will have to reject requests to deal with disputes that have already been submitted to another DSB or to a Member State's court or tribunal to prevent multiple parallel procedures.[129] Moreover, parties will need to pay to resort to these procedures. However, the Data Act does not specify the amount of the fees nor a way to calculate them. It merely provides for a transparency requirement, namely that the DSB 'shall make the fees, or the mechanisms used to determine the fees, known to the parties concerned before those parties request a decision'.[130] The procedure itself should be adversarial, as both parties should be able to express their point of view on the issue and to answer to the other party's submissions as well as to any potential expert statements.[131] These debates will nevertheless have to occur within a short time span, as the DSB will have to decide on the matter maximum ninety days after receiving the request.[132] It thus resembles an arbitration by experts mechanism.

One might question whether this ninety-days deadline is realistic. On the one hand, the parties may file lengthy submissions and require numerous expert statements. On the other hand, the DSB's decision will have to contain a statement of reasons supporting its findings, which means that it will have to 'answer' to the parties' (potentially lengthy) submissions and to the (potentially numerous) expert statements.[133] All of this could take a substantial amount of time.

Moreover, the DSB's decision, which must be delivered in writing or through another durable medium, will only be binding on the parties if they 'have explicitly consented to its binding nature prior to the start of the dispute settlement proceedings'.[134] Accordingly, an unwilling party (most likely a data holder unwilling to share data) can refuse to consent to the binding nature of this dispute mechanism. To be sure, a data recipient faced with an unwilling data holder retains its right to seek an effective remedy before Member States' courts or tribunals.[135] However, this will delay the dispute resolution, especially if the unwilling data holder drags its feet and is unwilling to cooperate. As a

128 Article 10.6 of the Data Act.
129 Article 10.7 of the Data Act.
130 Article 10.2 of the Data Act.
131 Article 10.8 of the Data Act.
132 Article 10.9 of the Data Act.
133 Ibid.
134 Articles 10.9 and 10.12 of the Data Act.
135 Article 10.13 of the Data Act.

result, while this speedy dispute resolution mechanism has been instituted in order to avoid lengthy discussions about the FRAND nature of the contractual terms, this will only work to the extent that both parties are willing to engage in such procedure and to agree to the binding nature of the decision. Otherwise, the dispute may have to be settled in court, which could take years, and this could hamper the development of innovative services by data recipients that need access to the data at hand.

In light of the above, it would have been welcome for the Data Act to make the DSB's decision binding on the parties.[136] Moreover, the Data Act could have made the dispute settlement mechanism a prerequisite to any FRAND-related procedure before national courts or tribunals. Indeed, the guarantees of independence, impartiality and expertise required to certify DSBs, as well as their duty to justify the reasons supporting their decision, should ensure the quality of their decisions, and should legitimate their binding nature. Moreover, providing that this speedy procedure is a prerequisite to a potential court case, this would have had the advantage of generating a first timely decision as to whether specific terms are FRAND. Overall, this could speed up the dispute resolution process and avoid unnecessary lengthy procedures that would hamper the core objective of the Data Act: enabling a wider use of data. It thus remains to be seen how effective the Data Act's dispute settlement mechanism will be in practice. More generally, this raises the question of appropriate enforcement, which we turn to now.

## 3. Enforcement

While the Data Act includes a chapter on implementation and enforcement (Chapter 9), the provisions leave quite some leeway and responsibility for Member States to set up the institutional frameworks in their territories. It should be noted that this discussion is to some extent separate from the discussion in section 2, which showed that private bargaining plays a large role in implementing the Data Act's baseline horizontal framework. At the same time, public enforcers remain ultimately responsible for interpreting and enforcing the different parts of the Data Act.

Two key aspects relating to enforcement will be discussed in this section: the competent authorities and the cross-border enforcement. Another relevant aspect that will not be covered here, but that is raised by others and therefore worth mentioning, is the uncertainty about whether private law remedies like injunctions and damages are available in case of breaches of the Data Act.[137]

136 See Podszun, R., & Offergeld, P. (2022). The EU data act and the access to secondary markets, *op. cit.*, p. 46.

137 Podszun, R., & Offergeld, P. (2022). The EU data act and the access to secondary markets, *op. cit.*, p. 45; Leistner, M., & Antoine, L. (2022). IPR and the use of open data and data sharing initiatives by public and private actors. *Study commissioned by the European Parliament's Policy Department for*

### 3.1 Competent Authorities

The Data Act requires Member States to designate one or more competent authorities as responsible for enforcement, and it gives Member States the choice to rely on existing authorities or establish new ones.[138] While the provisions leave a large margin of discretion to Member States to select competent authorities, a couple of preconditions are set by the Data Act. First, national data protection authorities are responsible for monitoring the application of the Data Act as far as the protection of personal data is involved.[139] Second, the competence of sectoral authorities has to be respected for specific sectoral data exchange issues relating to the implementation of the Data Act.[140] Third, the competent authority responsible for the enforcement of the provisions regarding the switching of data processing services must have experience in data and electronic communication services.[141] Within these boundaries, Member States are free to allocate responsibility for enforcement to one or more authorities. When a Member State designates more than one competent authority, the Member State has to select a data coordinator to facilitate cooperation among the different competent authorities.[142]

The lack of more prescriptive requirements regarding the selection of competent authorities at the national level has advantages and disadvantages. The main advantage is that Member States have the opportunity to choose the arrangement that best fits their national circumstances. For instance, depending on the resources and staffing of the various national authorities, Member States may prefer designating their data protection authority as the only competent authority for enforcing the Data Act. Alternatively, they may decide to divide responsibility across the data protection and competition authority and grant the coordinating role to the authority that still has the most available space to take up additional tasks. Beyond the availability of resources, the selection of the competent authorities may also stem from a policy choice at the national level. This points at the main disadvantage of the discretion of Member States to designate competent authorities, namely that harmonisation of enforcement is at risk. Even though the provisions of the Data Act regulate the conditions of data access, the exper-

*Citizens' Rights and Constitutional Affairs at the request of the Committee on Legal Affairs*, p. 118-119; Metzger, A., & Schweitzer, H. (2022). Shaping markets: A critical evaluation of the draft data act, *op. cit.*, p. 28-29; Drexl, J., Banda, C., Gonzalez Otero, B., Hoffmann, J., Kim, D., Kulhari, S.,... & Wiedemann, K. (2022). Position Statement of the Max Planck Institute for Innovation and Competition of 25 May 2022 on the Commission's Proposal of 23 February 2022 for a Regulation on Harmonised Rules on Fair Access to and Use of Data (Data Act), *op. cit.*, p. 89-90.

138 Article 37.1 of the Data Act.

139 Article 37.3 of the Data Act.

140 Article 37.4.a) of the Data Act.

141 Article 37.4.b) of the Data Act.

142 Article 37.2 of the Data Act.

tise of the respective authority may influence its attitude towards enforcement. A data protection authority will likely prioritise the protection of personal data in implementing the data access right, while a competition authority may let the need for sharing and reuse of data prevail. Even though the Data Act states that it is without prejudice to the protection of personal data,[143] there are always borderline cases where the expertise of the respective authority is likely to determine whether it leans more towards keeping datasets closed to prevent privacy concerns or opening datasets up to stimulate further innovation. Member States can thus steer implementation by selecting the data coordinator. On the one hand, the designation of different competent authorities among Member States may give rise to useful experimentation and help to sharpen implementation over time. On the other hand, effective coordination among Member States may be a challenge if the data coordinators have different fields of expertise and work with different vocabularies.

Different preferences have already been expressed regarding the designation of competent authorities. The EDPB and the EDPS have asked the EU legislator to designate national data protection authorities as data coordinators under the Data Act.[144] According to the EDPB and the EDPS, data protection authorities have 'a unique expertise, both legal and technical … placing them at the core of the digital regulation landscape'.[145] In their view:

> the designation of coordinating competent authorities other than data protection authorities could affect consistency in terms of monitoring the application of the provisions of the GDPR and lead to real complexity for digital players and data subjects.[146]

However, one may wonder whether data protection authorities have the most suitable expertise to implement the Data Act's data access right. The EDPB and the EDPS stress the importance of the fundamental right to the protection of personal data,[147] while the Data Act mainly stems from the need to create competitive and innovative data markets. As noted by Leistner and Antoine,[148] competition authorities may therefore be more suitable as data coordinators than data protection authorities. Competition authorities have relevant experience involving, for instance, setting the conditions of access and

143 Recital 7 of the Data Act.
144 EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal*, *op. cit.*, par. 113.
145 EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal*, *op. cit.*, par. 114.
146 EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal*, *op. cit.*, par. 116.
147 EDPB-EDPS, *Joint Opinion 2/2022 on the Data Act Proposal*, *op. cit.*, par. 115.
148 M. Leistner and L. Antoine, 'IPR and the use of open data and data sharing initiatives by public and private actors', *op. cit.*, p. 72.

the determination of FRAND terms. To ensure that the data access right is implemented in a data protection-compliant way but also in a way that fits the Data Act's objectives to promote competitive and innovative data markets, a mix of expertise is required. It may therefore be sub-optimal to leave all responsibility for enforcement with data protection authorities, who are already involved in the monitoring of the Data Act to ensure the protection of personal data.

Accordingly, we recommend Member States designate competition authorities as the data coordinators, who will then have to liaise with data protection authorities for aspects involving personal data. While the text of the Data Act also leaves Member States the choice to set up a new authority, we believe this would unnecessarily risk duplicating resources and make the enforcement landscape at the national level even more complicated. Because effective cooperation between competition, data protection or other authorities will be key to ensure proper implementation of the Data Act, we suggest Member States put cooperation protocols in place describing how the competent authorities can exchange insights and involve each other's expertise in monitoring the application of the Data Act. Not every Member State may currently have effective cooperation mechanisms in place.

Beyond this, we recommend aligning the enforcement of the Data Act and the Data Governance Act by making the same national authority responsible for coordinating enforcement.[149] There are synergies between the two Acts. In particular, the Data Act's data access right can boost the development of data intermediaries, services which facilitate data sharing between data holders and data users,[150] for which the Data Governance Act has established a notification framework that is governed at the national level.[151] We submit that consolidation of enforcement within the same national authorities is desirable to reap the benefits of the synergies between the two Acts and to ensure a harmonised implementation of data access-related legal mechanisms.

### 3.2 Cross-Border Enforcement

A large part of the data processing that falls within the scope of the Data Act is likely not confined to a single Member State. For instance, manufacturers will typically offer IoT devices to users in different Member States in parallel, and third parties wishing to access data under the Data Act may be based in a different Member State than the manufacturer or user of the IoT device. In such cross-border situations, the entity is

149 See also Colangelo, G. (2022). European Proposal for a Data Act–A First Assessment. *CERRE Evaluation Paper*, *op. cit*., p. 29.

150 Article 2.11 of the Data Governance Act.

151 Article 10-12 of the Data Governance Act.

subject to the competence of the Member State in which it has its main establishment.[152] The same approach applies in the GDPR, where the national data protection authority of the main establishment of the undertaking concerned is automatically responsible for acting as lead supervisory authority.[153] The latter approach has led to enforcement bottlenecks in EU data protection law because big tech firms in particular have their main establishments in countries like Ireland and Luxemburg that struggle to take up cross-border cases because they require larger resources than their populations allow for.[154] It is therefore remarkable that the Data Act makes the same choice.

One alternative would have been to rely on the approach of cross-border enforcement within EU consumer law, according to which the competent authorities select the national consumer authority that is best placed to coordinate the case.[155] This allows for more flexibility to allocate cases of cross-border relevance to national authorities. This kind of mechanism also ensures that the available resources are more evenly spread across cross-border issues than is currently the case for the enforcement of the GDPR. Interestingly, the Data Act does let national competent authorities allocate cases on a first-come-first-served basis when dealing with entities that do not have an establishment in the EU, until they have appointed a legal representative in one of the Member States.[156] It remains to be seen whether the Data Act's choice to make the Member State of the main establishment competent will lead to a similar enforcement bottleneck as under the GDPR. What is clear is that the legislator's choice allows for strategic behaviour, whereby entities can consider which authority will be responsible for supervising their activities for the purposes of the Data Act when moving their main establishment or designating their legal representative.

## 4. Conclusion

The Data Act aims to address issues that slow down the development of the European data economy by creating a legal instrument to enable wider data use across the economy. To do so, the Act creates a baseline horizontal framework for compulsory data

152 Article 37.10 of the Data Act.

153 Article 56.1 of the GDPR.

154 Communication from the Commission to the European Parliament and the Council, '*Data protection as a pillar of citizens' empowerment and the EU's approach to the digital transition – two years of application of the General Data Protection Regulation*', 24 June 2020, COM(2020) 264 final, p. 6; G. Gentile and O. Lynskey, 'Deficient by design? The transnational enforcement of the GDPR', *International & Comparative Law Quarterly*, 2022, Vol. 71, Issue 4, p. 799-830.

155 Article 17.2 of Regulation (EU) 2017/2394 of the European Parliament and of the Council of 12 December 2017 on cooperation between national authorities responsible for the enforcement of consumer protection laws (CPC Regulation), *OJ L 345/1*, 27 December 2017.

156 Article 37.11 and 13 of the Data Act.

sharing, which should provide incentives for horizontal data sharing across sectors. This is a truly novel approach compared to the previous disparate and mostly sector-specific data sharing approaches. Our analysis of the different provisions shows that it is possible to set conditions for data sharing that have a more general scope of application, but that there are several uncertainties in how they will be implemented. The success of the Data Act's horizontal framework for compulsory data sharing therefore largely depends on how its provisions are interpreted and applied. Several issues are worth keeping in mind in that process, including the need for guidance on what is 'reasonable' compensation for sharing data and on the interaction between data sharing obligations and the need to comply with personal data protection rules in situations where personal data have to be shared.

Despite the intention of the legislator to create clarity in the increasingly complex area of regulating data, it is important to highlight that the scope of the Data Act's horizontal framework is limited to compulsory B2B data sharing. This means that other legal frameworks will continue to exist for voluntary B2B data sharing, as well as for B2U, B2G, G2B and G2G data sharing. While this is due to the fact that different types of objectives are pursued depending on the type of data sharing, this might lead to an 'overload' of parallel data sharing regimes. This may hamper the fostering of data sharing more generally if there is no minimal alignment between the regimes.

The horizontal framework also only applies to the Data Act's IoT data access right and compulsory B2B data sharing obligations that enter into force after the date of application of the Data Act. Therefore, the horizontal framework does not apply to any pre-existing (sectoral) legislation that imposes compulsory B2B data sharing. In the short to medium term, this could lead to uncertainties for data holders due to the parallel application of different rules and regimes imposing compulsory B2B data sharing. It will thus be fundamental to ensure convergence between the Data Act's horizontal framework and these previously existing instruments in the coming years, possibly through anticipated revisions of these instruments if issues appear before their planned revision date. Similarly, as the Data Act provides that this is a baseline framework and that future legislation can go further in terms of data sharing requirements, it will be important to find the right balance between accommodating sector-specific needs on the one hand, and maintaining a minimum level of coherence between the different instruments on the other hand, in order to avoid a patchwork of data sharing regimes that would be hard to navigate.

Regarding enforcement, the Data Act leaves a lot of discretion to Member States to designate competent authorities and set up enforcement mechanisms. While the EDPB and the EDPS have advocated for designating national data protection authorities as coordinating competent authorities under the Data Act, competition authorities may possess more suitable expertise to take up this responsibility, considering the Data Act's

focus on promoting competitive and innovative data markets. To avoid the likely diversity of competent authorities designated by Member States from affecting the overall effectiveness of the enforcement system, we recommend Member States adopt cooperation protocols that describe how the different competent authorities can exchange insights and involve each other in monitoring compliance with the rules. Not every Member State will have this kind of cooperation protocol in place, while it is likely that authorities from different legal domains are involved in the application of the Data Act. Effective collaboration is thus vital to ensure proper implementation of the rules. Hopefully, the legislator's choice to make the Member State of the entity's main establishment competent for cross-border cases will still allow for a proper spread of responsibility across Member States and will not result in a similar enforcement bottleneck as in the GDPR.

To conclude, our analysis has illustrated the potential as well as the limits of the Data Act in creating a horizontal framework for data sharing. Its success in stimulating the data economy mainly depends on how ambiguities in the text are interpreted and what enforcement mechanisms are set up at the national level. As such, the Data Act forms an ambitious starting point for a next phase in the regulation of data sharing. While the Data Act's intention was to provide a clear and straightforward regime for data sharing, this chapter shows that many uncertainties are likely to remain in the coming years, at the very least until market players, regulators and courts start to find their way through the quagmire of different provisions and layers of regulation targeting data sharing in the EU.

PART

# IV

# RECONFIGURING THE LEGAL PARADIGM FOR THE PUBLIC SECTOR

CHAPTER

# IX

# Data-Driven Mergers in Healthcare: An EU Competition Law Perspective

**Tjaša Petročnik &**

**Inge Graef**[1]

https://doi.org/10.26116/hz58-jq20

1 PhD candidate and Associate Professor Tilburg Institute of Law, Technology, and the Society (TILT), Tilburg Law School (TLS), Tilburg University.

## 1. Introduction: Data and AI in Healthcare

Data has become a topic of special interest in various domains. One of these is healthcare, in which big data analytics and artificial intelligence (AI) have a great deal of potential.[2] Central to this potential is the idea that implementing AI solutions will revolutionise healthcare by yielding valuable insights into health, enabling the early detection and prediction of disease, more accurate diagnoses and more personalised clinical decision-making. Further, this will optimise workflows and increase efficiency, accelerate health research and empower patients, consequently improving health outcomes.[3] In recent years, AI systems have been applied in multiple domains of medicine, from detecting cancer to predicting infectious disease outbreaks and performing robotic surgeries.[4] Beyond the clinical context, a growing number of AI tools are offered directly to consumers, for instance apps that check for skin conditions and smartwatches that monitor blood pressure and other vital signs.[5]

The development of AI systems relies on the availability and linking of large amounts of heterogenous data to formulate problems, extract knowledge, and train algorithms, optimise them, and assess their performance. Algorithms also rely on data to provide predictions and input for decisions in the implementation stage.[6] In this sense, sharing

2 Dash, S., Shakyawar, S. K., Sharma, M., & Kaushik, S. (2019). Big data in healthcare: management, analysis and future prospects. *Journal of big data*, *6*(1), 1-25; Högberg, C., & Larsson, S. (2022). AI and Patients' Rights: Transparency and information flows as situated principles in public health care. In *De Lege–Yearbook Uppsala Faculty of Law 2021: Law, AI & Digitalization* (pp. 401-429). Iustus förlag.

3 Rinsche, F. (2017). *The Role of Digital Health Care Startups*, in Schmid, A. & Singh, S. (eds.), Crossing Borders – Innovation in the U.S. Health Care System (Schriften zur Gesundheitsökonomie, 84, Bayreuth: P.C.O.-Verlag 2017), 191; Pot, M., Kieusseyan, N., & Prainsack, B. (2021). Not all biases are bad: equitable and inequitable biases in machine learning and radiology. *Insights into imaging*, *12*(1), 13.; Sharon, T., & Lucivero, F. (2019). Introduction to the Special Theme: The expansion of the health data ecosystem–Rethinking data ethics and governance. *Big Data & Society*, *6*(2), 2053951719852969; Spatharou, A., Hieronimus, S., and Jenkins, J. (2020) *Transforming healthcare with AI: The impact on the workforce and organizations*. McKinsey & Company (10 March 2020) <https://www.mckinsey.com/industries/healthcare/our-insights/transforming-healthcare-with-ai>

4 Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC medicine*, *17*, 1-9; Morley, J., Machado, C. C., Burr, C., Cowls, J., Joshi, I., Taddeo, M., & Floridi, L. (2020). The ethics of AI in health care: a mapping review. *Social Science & Medicine*, *260*, 113172; Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S.,... & Schönlieb, C. B. (2021). Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans. *Nature Machine Intelligence*, *3*(3), 199-217; Van Kolfschooten, H. (2022). EU regulation of artificial intelligence: Challenges for patients' rights. *Common Market Law Review*, *59*(1), 81-112.

5 See, e.g., Babic, B., Gerke, S., Evgeniou, T., & Cohen, I. G. (2021). Direct-to-consumer medical machine learning and artificial intelligence applications. *Nature Machine Intelligence*, *3*(4), 283-287; Gerke, S., & Rezaeikhonakdar, D. (2022). Privacy aspects of direct-to-consumer artificial intelligence/machine learning health apps. *Intelligence-Based Medicine*, *6*, 100061.

6 Bohr, A., & Memarzadeh, K. (2020). The rise of artificial intelligence in healthcare applications. In

and using data is central to AI and its benefits; it is thus also an important policy objective in healthcare,[7] believed to contribute to longer and healthier lives.[8] Further, it is considered a means to support 'a thriving European data economy'[9] where 'access to data and the ability to use it are essential for innovation and growth.'[10] This includes healthcare and AI's potential in it.[11]

The growing role of data and AI makes healthcare a relevant sector to study from the perspective of data-centric regulatory interventions. This chapter takes an EU competition law perspective on AI and data-driven mergers in healthcare. It analyses two recent decisions in this domain that garnered public attention: the Google/Fitbit and the Illumina/Grail cases. The commercialisation of healthcare has led to prominent mergers and acquisitions that raise questions about how EU merger control understands and deals with data, and if and how it interacts with non-economic considerations. While having access to large amounts of data is necessary to develop and implement AI and other data-driven technologies, sharing and using this data can pose several challenges.[12] This includes market concentration problems,[13] but also problems that are not primarily economic in nature, like diminished privacy, opaque decision-making and a lack of explainability,[14] quality and safety concerns related to healthcare AI,[15] and biased outputs

*Bohr, A., & Memarzadeh, K. (Eds.). (2020). Artificial intelligence in healthcare. Academic Press*. Högberg and Larsson, *supra* n. 2, 404, 421; Tohka, J., & Van Gils, M. (2021). Evaluation of machine learning algorithms for health and wellness applications: A tutorial. *Computers in Biology and Medicine*, *132*, 104324; van Kolfschooten, *supra* n. 4.

7 Riso, B., Tupasela, A., Vears, D. F., Felzmann, H., Cockbain, J., Loi, M.,... & Rakic, V. (2017). Ethical sharing of health data in online platforms–which values should be considered?. *Life sciences, society and policy*, *13*, 1-27.

8 See, e.g., *Communication From The Commission To The European Parliament, The Council, The European Economic And Social Committee And The Committee Of The Regions A European strategy for data* COM(2020) 66 final.

9 Marelli, L., Lievevrouw, E., & Van Hoyweghen, I. (2020). Fit for purpose? The GDPR and the governance of European digital health. *Policy studies*, *41*(5), 447-467.

10 *European data strategy*, European Commission (2024), <https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en>.

11 Schneider, G. (2020). Health data pools under European policy and data protection law: Research as a new efficiency defence?. *J. Intell. Prop. Info. Tech. & Elec. Com. L.*, *11*(1), 49-67; van Kolfschooten, *supra* n. 4.

12 Högberg and Larsson, *supra* n. 2, 421.

13 Massolo, A. (2020). *Mergers in Big Tech: An overview of EU and national case law*, Concurrences (18 August 2022), <https://www.concurrences.com/en/bulletin/special-issues/mergers-in-big-tech/mergers-in-big-tech-an-overview-of-eu-and-national-case-law>.

14 Carusi, A., Winter, P. D., Armstrong, I., Ciravegna, F., Kiely, D. G., Lawrie, A.,... & Swift, A. (2023). Medical artificial intelligence is as much social as it is technological. *Nature Machine Intelligence*, *5*(2), 98-100; van Kolfschooten, *supra* n. 4.

15 Akbar, S., Coiera, E., & Magrabi, F. (2020). Safety concerns with consumer-facing mobile health applications and their consequences: a scoping review. *Journal of the American Medical Informatics Association*, *27*(2), 330-340.

and discrimination.[16] Altogether, these developments can potentially affect access to healthcare and exacerbate health inequalities.[17]

Against this background, this chapter addresses whether and how merger assessments can give adequate weight to non-economic interests in data-driven mergers in healthcare. While the application of EU competition law is mainly based on economic considerations, economic and non-economic interests are becoming increasingly intertwined in data-driven healthcare. As such, our approach is mainly grounded in the disciplinary perspective of law and technology that approaches technological innovation with a view to ensure compliance with norms and values and the realisation of expected benefits while minimising potential harms.[18]

The chapter is structured as follows: after a more general introduction to the role and value of data in healthcare, a look into the data economy more broadly in section 2, and an overview of the Google/Fitbit and Illumina/Grail merger decisions in section 3, section 4 offers recommendations for incorporating non-economic interests into the assessment of data-driven healthcare mergers more proactively within as well as beyond EU merger review.

## 2. Understanding Health Data[19]

Generally, health data can be understood as 'any type of data that is useful for improved research, innovation and healthcare-related decision making', stemming from diverse sources ranging from patient records, data from registries, health research data, data from digital devices (e.g. wearables and smartphones) and environmental and other sensors, and is used for a broad variety of purposes.[20] As data can be in several places

16 Jabłonowska, A., Kuziemski, M., Nowak, A. M., Micklitz, H. W., Pałka, P., & Sartor, G. (2018). Consumer law and artificial intelligence. EUI Department of Law Research Paper No. 2018/11, <https://cadmus.eui.eu/handle/1814/57484>; Kooli, C., & Al Muftah, H. (2022). Artificial intelligence in healthcare: a comprehensive review of its ethical concerns. *Technological Sustainability*, *1*(2), 121-131. See also *Google's new AI skincare tool may not work on patients with darker skin tones*, Euronews (26 May 2021), <https://www.euronews.com/next/2021/05/26/google-s-new-ai-skincare-tool-may-not-work-on-patients-with-darker-skin-tones>.

17 Stavroulaki, T. (2022). Mergers That Harm Our Health. *Berkeley Business Law Journal*, *19*, 89. Stavroulaki, T. (2021) *Opening the Black Box: The Hidden Costs of Data-Driven Mergers in Health Care*, Promarket (14 May 2021), <https://www.promarket.org/2021/05/14/costs-data-driven-mergers-health-care-insurers-drug-suppliers/>.

18 See Butenko, A., & Larouche, P. (2015). Regulation for innovativeness or regulation of innovation?. *Law, Innovation and Technology*, *7*(1), 52-82.

19 The discussion in section 2 is partly based on the forthcoming PhD manuscript of Tjaša Petročnik.

20 These purposes range from personal health management to medical interventions, health policy and planning, and business competitiveness. Further, insurers might be interested in data to control costs, hospitals to optimise clinical workflows, and pharmaceutical companies might want to rely on it as a basis for drug discovery. Marjanovic, S., Ghiga, I., Yang, M., & Knack, A. (2018).

simultaneously and has the ability to be aggregated, used in conjunction or reused within the same or in other contexts,[21] harnessing and making sense of this data is possible 'as never before' in light of new technologies.[22] Still, due to its often personal and sensitive nature, access to and the use of certain kinds of health data needs to be restricted.[23] This can limit or challenge data sharing aspirations due to ethical and other concerns.

To characterise data sharing and its role, an array of metaphors have been put forward in the literature.[24] For the purposes of this paper, we conceptualise health data in terms of several dimensions. Data, particularly personal data, has been considered to have a strong link to individuals,[25] and it has thus been conceived of as an expression or manifestation of the self,[26] or even something constitutive of an individual.[27] This reflects a somewhat dignitarian aspect to health data, which is often related to the most intimate aspects of people's lives. On the other hand, data has been viewed as an input or resource for other goods or practices, having value derived from its use or utility.[28] This duality

Understanding value in health data ecosystems: A review of current evidence and ways forward. *Rand health quarterly*, *7*(2). Petrocnik, T. (2022). Health data between improving health (care) and fuelling the data economy. *Technology and Regulation*, *2022*, 124-127. Sharon and Lucivero, *supra* n. 3.

21 McMahon, A., Buyx, A., & Prainsack, B. (2020). Big data governance needs more collective responsibility: the role of harm mitigation in the governance of data use in medicine and beyond. *Medical law review*, *28*(1), 155-182. As pointed out by Prainsack, it is increasingly harder in this context to establish who can access, control, and use data. See Prainsack, B. (2019). Logged out: Ownership, exclusion and public value in the digital data and information commons. *Big Data & Society*, *6*(1), 2053951719829773.

22 Fry, E., & Mukherjee, S. (2018). Big data meets biology. *Fortune*, *4*, 24-35. <https://fortune.com/2018/03/19/big-data-digital-health-tech/>.

23 *The role of data in AI business models (report)*, Open Data Institute (20 April 2018), 11, <https://theodi.org/insights/reports/the-role-of-data-in-ai-business-models/>.

24 These liken health data to everything from the new oil to the new blood. Perakslis, E., & Coravos, A. (2019). Is health-care data the new blood?. *The Lancet Digital Health*, *1*(1), e8-e9. Stevens, M., Wehrens, R., & De Bont, A. (2018). Conceptualizations of Big Data and their epistemological claims in healthcare: A discourse analysis. *Big Data & Society*, *5*(2), 2053951718816727.

25 Jurcys, P., Donewald, C., Fenwick, M., Lampinen, M., & Smaliukas, A. (2020). Ownership of user-held data: Why property law is the right approach. *Harvard Journal of Law and Technology Digest [2021]*, referring to personality theory. Personality rights, in this context, are internal and inalienable. See Hummel, P., Braun, M., & Dabrock, P. (2021). Own data? Ethical reflections on data ownership. *Philosophy & Technology*, *34*(3), 545-572. See further Van der Sloot, B. (2015). Privacy as personality right: Why the ECtHR's focus on ulterior interests might prove indispensable in the age of big data. *Utrecht Journal of International and European Law*,*31*, 25.

26 Jurcys and others, *supra* n. 25, 22.

27 Floridi, L. (2016). On human dignity as a foundation for the right to privacy. *Philosophy & Technology*, *29*, 307-312.

28 Tombal, T. (2020). Economic dependence and data access. *IIC-International Review of Intellectual Property and Competition Law*, *51*(1), 70-98. Graef, I., Gellert, R., & Husovec, M. (2018). Towards a holistic regulatory approach for the European data economy: Why the illusive notion of non-personal data is counterproductive to data innovation. TILEC Discussion Paper No. 2018-029 (2018),

is increasingly evident in the EU data governance regime, which aims to protect individuals with regard to the processing of personal data while at the same time promoting its free flow[29] for diverse purposes. In this light, we can understand data in the context of a multiplicity of goals and policy objectives, manifesting on the individual, private or socio-economic level. To elaborate, health data is linked to individuals' private interests, like self-care, identity, and informational self-determination. At the same time, it can contribute to knowledge production, such as in regard to health policy and research, and improving the accuracy, efficiency and accessibility of healthcare systems,[30] which face increased costs and demands, administrative burdens and staff shortages.[31] Data is also central to facilitating innovation and fuelling the digital economy with many businesses endeavouring to deliver upon the promises of AI.[32]

### 2.1 Data-Driven Business Models

With regard to the data economy, AI forms the foundation of a range of different products and services in healthcare. This leads firms to adopt data-driven business models[33] and to depart on a 'new gold rush' to obtain, analyse and leverage data[34] as a central competitive parameter on digital markets.[35] It is increasingly being acknowledged that data can result in competitive advantages,[36] although 'data doesn't automatically equal power'.[37] Importantly, it is not necessarily only the individuals' data that is valuable; the value of data lies in aggregated information used for profiling and relating people 'to

<https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3256189>. For a similar distinction between dignitarian and consequentialist views on data collection and use, see Jablonowska and others, *supra* n. 16.

29 *European data strategy*, European Commission (2024), https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en.

30 Van Kolfschooten, *supra* n. 4, 9; see also Stevens, Wehrens, and De Bont, *supra* n. 24; Riso and others, *supra* n. 7, 12.

31 Högberg and Larsson, *supra* n. 2, 404; Rinsche, *supra* n. 3, 190.

32 Perakslis and Coravos, *supra* n.24; see also Petročnik, *supra* n. 20, 125; Rinsche, *supra* n. 3, 187-8.

33 Stucke, M. E. & Grunes, A. P. (2016). Big Data and Competition Policy. Oxford University Press; see also Open Data Institute, *supra* n. 23, 10-1.

34 Fry and Mukherjee, *supra* n. 22.

35 Majcher, K. & Robertson, V. H.S.E. (2022), The Twin Transition to a Green and Digital Economy: The Role for EU Competition Law (May 11, 2022). Graz Law Working Paper No. 05-2022, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4106485>; see also Allen P. Grunes and Maurice E. Stucke, No Mistake About It: The Important Role of Antitrust in the Era of Big Data, University of Tennessee Legal Studies Research Paper No. 269 (2015), <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2600051>.

36 *Privacy and competitiveness in the age of big data*, European Data Protection Supervisor (26 March 2014), <https://www.edps.europa.eu/press-publications/press-news/press-releases/2014/privacy-and-competitiveness-age-big-data_en>.

37 Sayer, P. (2016) *Big data is an antitrust issue too, says European Commissioner*, PCWorld (18 January 2016),

one another on the basis of relevant shared population features',[38] and thus in predictions and its future utility.

In this sense, as the value of data is hard to establish when it is first created or collected,[39] innovation on digital markets is driven by ways to secure control of data as an asset and the key resource of future business success.[40] Obtaining data to capture its value can be achieved by various means, such as by expanding and engaging with businesses' own customer base, running a two-sided platform, partnership agreements, purchasing data from data brokers or through merger activity.[41] The latter, as the strategy to acquire and use data as a competitive advantage in data-driven business models,[42] is the focus of the present chapter.

## 2.2 The Harms of Data-Driven Business Strategies

Before turning our attention to the regulation of merger activity, a discussion on the possible harms that might stem from business strategies focused on acquiring and using data and AI is necessary. We understand these broadly, ranging from harms to individual and group interests to harms to societal interests,[43] like harms to competition, but also harms to other public interest considerations.[44] We suggest that, just as the under-

<https://www.pcworld.com/article/419229/big-data-is-an-antitrust-issue-too-says-european-commissioner.html>.

38 Viljoen, S. (2021). A relational theory of data governance. *Yale LJ*, *131*, 573; Caffarra, C., & Valletti, T. (2020). *Google/Fitbit review: Privacy IS a competition issue*. Vox EU/CEPR. March, 4. <https://cepr.org/voxeu/blogs-and-reviews/googlefitbit-review-privacy-competition-issue>. See also Carmi, E. (2021). A feminist Critique to digital consent. *Seminar.net*, 17(2). <https://doi.org/10.7577/seminar.4291>.

39 Barocas and Nissenbaum in Carmi, *supra* n. 38.

40 Birch, K., Chiappetta, M., & Artyushina, A. (2020). The problem of innovation in technoscientific capitalism: data rentiership and the policy implications of turning personal digital data into a private asset. *Policy studies*, *41*(5), 468-487; see also Grunes and Stucke, *supra* n. 35, 8.

41 Chen, Z., Choe, C., Cong, J., & Matsushima, N. (2022). Data-driven mergers and personalization. *The RAND Journal of Economics*, *53*(1), 3-31. Grunes and Stucke, *supra* n. 35, 7; European Data Protection Supervisor, *supra* n. 36, 10-1; Lynskey, O. (2019). Grappling with "data power": normative nudges from data protection and privacy. *Theoretical Inquiries in Law*, *20*(1), 189-220. Note that a merger is defined as 'amalgamation or joining of two or more firms into an existing firm or to form a new firm'. Mwemba, W. (2024) *Merger (notion)*, Concurrences (2024), <https://www.concurrences.com/en/dictionary/Concentration>.

42 Grunes and Stucke, *supra* n. 35, 3.

43 Smuha, N. A. (2021). Beyond the individual: governing AI's societal harm. *Internet Policy Review*, *10*(3); Graef, I., & van der Sloot, B. (2022). Collective data harms at the crossroads of data protection and competition law: Moving beyond individual empowerment. *European Business Law Review*, *33*(4), 513-536.

44 Binns, R., & Bietti, E. (2020). Dissolving privacy, one merger at a time: Competition, data and third party tracking. *Computer Law & Security Review*, *36*, 105369.

standing of data reflects the multiplicity of various aspirations, so should our conceptualisation of harms. This reflects the insight that data is made useful and thus valuable to a certain actor for a certain aim.[45] To keep the analysis manageable, the typology we propose here differentiates between economic and non-economic data-related harms. Broadly, this also corresponds to the parallel health policy goals of efficiency and equity[46] respectively, and arguably to the identified purposes of data sharing and use too. First, we identify data-related harms that are predominantly economic in nature and linked to efficiency. On the one hand, data availability may provide consumer benefits in the form of higher quality, personalised products and services based on data that can address unmet needs with greater efficiency and speed. On the other hand, it might lead to market concentration or exclusion, as data-driven business models presuppose that a firm will need to maintain the data advantage over its rivals.[47] Data accumulation might also lead to exploitation if individuals' preferences and decisions are steered in directions they would not necessarily have gone otherwise,[48] or if used for the identification and discrimination of higher-risk consumers.[49]

Second, the non-economic data-related harms are relevant. Namely, while economic efficiency is a coveted goal in modern healthcare, there are other equally important objectives, including equity considerations.[50] The literature notes that these might also be affected by data-driven business strategies that aim to acquire and use data. Concretely, these harms can pertain to violations of privacy and data protection, and transparency and information provision that are particularly relevant in healthcare contexts.[51]

45 Helmond, A. and van der Vlist, F. (2023) Situating the Marketization of Data. In K. van Es and Verhoeff, N. (Eds.), *Situating Data. Inquiries in Algorithmic Culture*. Amsterdam University Press.

46 This distinction is suggested by Sauter, W. (2012). The impact of EU competition law on national healthcare systems. TILEC Discussion Paper No. 2012-032, <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2138175>.

47 Grunes and Stucke, *supra* n. 35, 10. See further Salzberger, D., Iatrou, N., Kwinter, G., & Keogh, E. (2021). Data, Not Data: Uncovering the Implications of Data in Merger Reviews. *U. Mem. L. Rev.*, *52*, 969.

48 Graef and van der Sloot, *supra* n. 43, 515.

49 As noted, while the overall welfare might still increase even if certain consumers would e.g. be required to pay more, leading to a desired outcome from the efficiency perspective, this might not correspond to healthcare goals that are non-economic in nature and could reinforce or even increase health disparities. Stavroulaki (2022), *supra* n. 17, 95-9, 108.

50 See Sauter, *supra* n. 46.

51 Van Kolfschooten, *supra* n. 4. *EU Commission assessment of the Google-Fitbit merger must include human rights risks*, Amnesty (27 November 2020), <https://www. amnesty.eu/news/eu-commission-assessment-of-the-google-fitbit-merger-must-include-human-rights-risks>. Here it is necessary to note that while privacy can be considered an economic parameter of competition, as e.g. in the Facebook/WhatsApp merger, for present purposes we consider it a non-economic data-related harm as the scope of privacy is not limited to economic considerations.

Further, there are concerns linked to the reliability and accuracy of data and AI[52] and doubts regarding AI's performance and utility in practice, in both clinical and consumer contexts.[53] Non-economic data-related harms can in this vein manifest in the over- or under-diagnosing of certain conditions. This can result in unnecessary medical procedures and thus the unnecessary use of resources or, conversely, missing of an actual health condition.[54] Further, there is a concern that AI systems might perform better for sub-populations that are better represented in the datasets, while providing less accurate results for groups and contexts for which the data is under-inclusive.[55] This can lead to or exacerbate biases in healthcare, health inequalities and discrimination.[56] On a more fundamental level, some have expressed apprehension that the move towards personalised healthcare enabled by AI and data analytics can ultimately challenge the solidarity underpinning healthcare systems.[57]

## 3. An Overview of the Role of Data and AI in the EU's Approach to Mergers

The healthcare sector is no stranger to mergers,[58] including those that can primarily be considered as data-driven.[59] While merger activity can lead to efficiencies and other beneficial consequences for consumers and society more broadly, some mergers might

52 Akbar, Coiera, and Magrabi, *supra* n. 15; Babic and others, *supra* n. 5; Kelly and others, *supra* n. 4; Wolf, J. A., Moreau, J. F., Akilov, O., Patton, T., English, J. C., Ho, J., & Ferris, L. K. (2013). Diagnostic inaccuracy of smartphone applications for melanoma detection. *JAMA dermatology*, *149*(4), 422-426.

53 Tung, J. Y., Shaw, R. J., Hagenkord, J. M., Hackmann, M., Muller, M., Beachy, S. H.,... & Ginsburg, G. S. (2018). Accelerating precision health by applying the lessons learned from direct-to-consumer genomics to digital health technologies. NAM Perspectives. Discussion Paper, National Academy of Medicine, Washington, DC. <https://doi.org/10.31478/201803c>. Heaven, W.D (2020) *Google's medical AI was super accurate in a lab. Real life was a different story*, MIT Technology Review (27 April 2020) <https://www.technologyreview.com/2020/04/27/1000658/google-medical-ai-accurate-lab-real-life-clinic-covid-diabetes-retina-disease>.

54 Abbott, L. M., & Smith, S. D. (2018). Smartphone apps for skin cancer diagnosis: Implications for patients and practitioners. *Australasian Journal of Dermatology*, *59*(3), 168-170; Akbar, Coiera, and Magrabi, *supra* n. 15, 330-40; Babic and others, *supra* n. 5, 283-7; Wolf and others, *supra* n. 52, 422-6.

55 Kelly and others, *supra* n. 4.

56 *Ethics guidelines for trustworthy AI*, High-Level Expert Group on AI (8 April 2019), <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>.

57 McFall, L. (2019). Personalizing solidarity? The role of self-tracking in health insurance pricing. *Economy and society*, *48*(1), 52-76; Prainsack, B. (2018). The "we" in the "me" solidarity and health care in the era of personalized medicine. *Science, Technology, & Human Values*, *43*(1), 21-44.

58 See, e.g., Van De Gronden, J., & Szyszczak, E. (2014). Introducing competition principles into health care through EU law and policy: a case study of the Netherlands. *Medical Law Review*, *22*(2), 238-254.

59 See, e.g., Chen and others, *supra* n. 41; see also Sheehy Jr, T. (2019). The Consumer Healthcare Data Market: Redefining Healthcare Mergers through the Linked Consumer Retail Data and Health Insurance Markets. *University of Baltimore Law Forum*, *50*(1), 28-56.

have undesirable effects on the market. To safeguard against these, a merger control system has been established in many jurisdictions, including the EU.[60]

### 3.1 EU Merger Control: A Short Introduction

The purpose of the merger control in the EU is to look 'to the future' to prevent mergers that may reduce competition in a market, usually by the creation or strengthening of dominance, and thus maintain competitive markets.[61] Regulation (EC) No 139/2004 on the control of concentrations between undertakings (the EU Merger Regulation) represents the EU regulatory framework and is accompanied by several notices and guidelines. According to the EU Merger Regulation, mergers that involve market activity in several Member States and exceed certain turnover thresholds (i.e. that have a Union dimension) have to be notified to and examined on the EU level[62] by the European Commission Directorate General for Competition before consummation.[63]

The overall objective of EU merger control is to assess whether the merger would 'significantly impede effective competition' (the 'SIEC test') due to, but not exclusively, the creation or strengthening of a dominant position on the market.[64] If the merger significantly impedes effective competition, it is to be considered incompatible with the internal market; in this sense, 'the leading litmus test' since the EU Merger Regulation was introduced is significant harm to competition, rather than dominance.[65] Efficiencies can be taken into account when determining the impact of a merger on the competition in the market, but it is primarily economic benefits that are considered as a justification.[66] Following an investigation – a procedure with strict deadlines in which first the relevant market is defined, then a competitive assessment performed[67] – the European

60 Whish, R., & Bailey, D. (2021). *Competition law*. Oxford University Press.

61 Whish and Bailey, *supra* n. 60, 828-37; *Competition Policy – Mergers*, European Commission (2024), <https://competition-policy.ec.europa.eu/mergers_en>.

62 See Articles 1 and 2 of the Council Regulation (EC) No 139/2004 on the control of concentrations between undertakings (2004) OJ L 24 (EU Merger Regulation).

63 The EU Merger Regulation also foresees that the Commission refers a notified merger to the competent authorities in a Member State, where the merger significantly affects competition in a national market, or when such a request is made. See Articles 4(4) and 9 of the EU Merger Regulation. Furthermore, a merger without the Union dimension can be referred by Member State(s) to the Commission. See Articles 4(5) and 22 of the EU Merger Regulation.

64 See Article 2(2) and (3) of the EU Merger Regulation.

65 Monti, G. (2022). The merger control regulation: What's past is prologue. *Common Market Law Review*, *59* (SI), 13-24.

66 See Recital 29 of the EU Merger Regulation; see also Hancher, L., & Sauter, W. (2012). *EU competition and internal market law in the health care sector*. Oxford University Press. See further Monti, *supra* n. 65, 19.

67 See Article 6 of the EU Merger Regulation.

Commission can decide to allow or prohibit the merger, or to allow it subject to certain conditions[68] (e.g. divestiture)[69] to remedy the competition problems identified in the assessment.[70]

### 3.2 EU Merger Control in Relation to Data and Digital Markets

Access to data is a common potential competition concern in the digital context that is characterised by first-mover advantage and an industry dynamic that favours concentration.[71] There is a risk, however, that relevant acquisitions occasionally fly under the radar of EU merger control, despite the concerns that they are motivated by the goal of acquiring data and that companies might engage in 'killer acquisitions' or raise entry barriers. This is because firms' turnover may remain below established jurisdictional thresholds.[72] As argued, turnover metrics can be problematic given that many firms in the digital context, which are the target of acquisitions, provide their products or services free of charge, thereby generating low revenues while retaining economic value in terms of user knowledge, data and network effects.[73]

In digital mergers that have nonetheless been assessed thus far, the Commission engaged with data-related theories of harm – in addition to other possible competitive potential concerns – in the following manner. It looked into the accumulation of data[74] and potential input foreclosure issues.[75] To illustrate, in the Facebook/WhatsApp merger, the Commission among other things addressed competition concerns related to Facebook using data from WhatsApp users to strengthen its position in advertising. It ultimately concluded that, even if this was the case,[76] rivals would have sufficient data to

68 See Article 8 of the EU Merger Regulation.

69 See Commission notice on remedies acceptable under Council Regulation (EC) No 139/2004 and under Commission Regulation (EC) No 802/2004.

70 Whish and Bailey, *supra* n. 60, 842, 848-9, 907.

71 Robertson, V. H. S. E. (2022). Merger review in digital and technology markets: Insights from national case law. <https://competition-policy.ec.europa.eu/system/files/2022-12/kd0422317enn_merger_review_in_digital_and_tech_markets_1.pdf>; Grunes and Stucke, *supra* n. 35, 8.

72 For instance, both the Facebook/WhatsApp and Meta/Kustomer acquisitions did not meet the turnover thresholds and were only assessed by the European Commission after referral by the merging parties, in the former case, and national competition authorities, in the latter case. See Case No COMP/M.7217 – *Facebook/WhatsApp*, 3 October 2014, paras 9-12 and Case M. 10262 – *Meta (formerly Facebook)/Kustomer*, 27 January 2022, paras 5-6.

73 Borgogno, O., & Zangrandi, M. S. (2022). Data governance: a tale of three subjects. *Journal of Law, Market & Innovation*, *1*(2), 50-75; Robertson, *supra* n. 71, 26; see also Lynskey, *supra* n. 41, 215.

74 Case M. 10262 – *Meta (formerly Facebook)/Kustomer*, 27 January 2022, paras 560-1

75 Case M.8124 – *Microsoft/LinkedIn*, 6 December 2016, paras 246-77.

76 The Commission did not account for cross-platform data aggregation, as this would supposedly go against WhatsApp's privacy policy and would require overcoming technical obstacles to do so. Case M.7217 – *Facebook/WhatsApp*, para 185. Later, the Commission fined Facebook for providing

compete as 'there will continue to be a large amount of Internet user data that are valuable for advertising purposes and that are not within Facebook's exclusive control'.[77] Similarly, in the Google/DoubleClick case, the Commission looked into data combinations. It reasoned that a possible data combination would unlikely squeeze out competitors, as similar web-usage data is readily available to Google's competitors.[78] These predominantly economic aspects of data stem from its role as an input or an asset and can manifest in increasing or leveraging market power or entry barriers for competitors.[79]

The Commission's decisional practice also touched on other aspects, which considered consumers' interests as well as other interests in terms of data. With regard to consumers' personal data, the Commission explicitly stated that privacy-related concerns that might emerge due to data concentration or use 'do not fall within the scope of the EU competition law rules but within the scope of the EU data protection rules'.[80] Nevertheless, data privacy has been brought into competitive analyses. For instance, it was considered as an aspect of monetisation,[81] an element of product quality,[82] consumer choice[83] and innovation.[84] While privacy has been considered in the assessment of mergers when it is a parameter on the basis of which the market players compete (for instance replacing price as the main traditional parameter of competition), the Commission has thus

misleading information in this respect. Case M.8228 – *Facebook/WhatsApp*, 18 May 2017. See further, e.g., Lynskey, *supra* n. 41, 216-7, pointing to possible negative impacts of cross-platform data aggregation on users' privacy.

77 Case M.7217 – *Facebook/WhatsApp*, paras 187-9.

78 Case No COMP/M.4731 – *Google/DoubleClick*, 11 March 2008, paras 364-6. See further Chirita, A. D. (2020) *Data-Driven Mergers under EU Competition Law*, in Akseli, O., & Linarelli, J. (Eds.). (2020). *The Future of Commercial Law: Ways Forward for Change and Reform* (Vol. 4). Bloomsbury Publishing; Deutscher, E. (2017) How to Measure Privacy-Related Consumer Harm in Merger Analysis? A Critical Reassessment of the EU Commission's Merger Control in Data-Driven Markets. Faculty of Law, Stockholm University Research Paper No. 40 <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3075200>; Beadouin, Y. et al. (2022), *Merger Enforcement in Digital and The Markets: an Overview of the European Commission's Practice*, Competition Policy Brief n. 2.

79 Data Protection Supervisor, *supra* n. 36, 30.

80 Case M.7217 – *Facebook/WhatsApp*, para 164; Case M.7813 – *Sanofi/Google/DMI JV*, 23 February 2016, para 70.

81 See, e.g., Case M.7217 – *Facebook/WhatsApp*.

82 Case M.7217 – *Facebook/WhatsApp*, paras 87-90; Case No COMP/M.5727 – *Microsoft/Yahoo! Search Business*, 18 February 2010, para 204. See also Case No COMP/M.4854, *TomTom/TeleAtlas*, 14 May 2008, paras 272-5, where client confidentiality concerns were seen as similar to product degradation.

83 Case M.8124 – *Microsoft/LinkedIn*, para 350.

84 Case M.8124 – *Microsoft/LinkedIn*, para 246; Case M.7217 – *Facebook/WhatsApp*, para 87.

far not taken into account the impact mergers can have on privacy as a non-economic consideration directly.

The increasing relevance of data in healthcare has also already featured in merger cases. A prominent example is the Google/Sanofi case, addressing a joint venture set up by the two parties to offer services for the management and treatment of diabetes using an integrated digital e-medicine platform. One of the concerns raised in the case was whether the joint venture would be able to lock patients into its services by restricting their ability to transfer data to other platforms. Since the GDPR provides individuals with a right to data portability, the Commission took the view that the merging parties would lack the ability to lock in patients.[85] Irrespective of the adequateness of this reasoning,[86] the case illustrates how individuals' dependence, in this case diabetes patients, links economic and non-economic interests. Locking in patients can reduce competition in healthcare markets, but it may also affect their ability to access healthcare services that align with their choice or needs, thus affecting the quality of healthcare they receive. While the Commission did not consider this link in Google/Sanofi, merger reviews will increasingly encounter situations where economic and non-economic interests overlap.[87]

85 Case M.7813 – *Sanofi/Google/DMI JV*, paras 67-69.

86 Criticising this reasoning because it presumes competition concerns are absent when other regimes have relevant legal requirements in place, see Podszun, R. (2018). *Dismembering producers from customers: The Google/Sanofi Joint Venture*. CPI Antitrust Chronicle, (February 2018), <https://www.competitionpolicyinternational.com/wp-content/uploads/2018/02/CPI-Podszun.pdf>.

87 Some other merger cases in the field of healthcare data-driven technologies in the past decade include: Case No COMP/M.7337 – *IMS Health/Cegedim Business*, 19 December 2014, which concerned tools for customer relationship management, marketing, and data management for healthcare industries. The Commission approved the merger, subject to conditions to divest parts of its primary market research business and to allow third party access to its sales tracking data structure; Case M.8061 – *IMS Health/Quintiles*, 12 August 2016, concerning a merged entity that would provide services to medicine and medical devices manufacturers to support running clinical trials or track sales. The Commission did not oppose the merger; Case M.8991 – *Alphabet/Resmed/JV*, 1 October 2018, establishing a joint venture that aimed to study health and financial impacts of untreated sleep apnoea and other similar sleep disorders and develop software solutions for their more efficient diagnosis and treatment. The Commission did not oppose the operation; Case M.9812 – *Verily Life Sciences/Santen Pharmaceutical/JV*, 3 August 2020, established to study and develop ophthalmology devices and digital technologies to diagnose or treat eye disorders. The Commission did not oppose the operation; and Case M.9945 – *Siemens Healthineers/Varian Medical Systems*, 19 February 2021, concerning medical imaging and radiotherapy solutions. The merger was approved, subject to conditions to respect interoperability standards in the sector.

## 3.3 The Commission's Approach in the Google/Fitbit and Illumina/Grail Merger Cases

In this section, we analyse two recent merger cases in the health domain: namely the Google/Fitbit and Illumina/Grail mergers. We put particular focus on how the Commission's decisions engaged with data-driven technologies and AI, and how they tackled the non-economic harms in the proposed theories of harm, if at all. The two cases are noteworthy on two points. First, they are relevant for the assessment of economic considerations of data as an asset, in the context of two non-horizontal mergers. These types of mergers are not usually considered problematic from the perspective of competition law, since the merging entities are not rivals, and are even considered to increase consumer welfare.[88] Next, the cases are interesting from the perspective of emerging paradigms in healthcare and the various health policy goals they promote. Specifically, the Google/Fitbit merger revolves around harnessing wearable technologies, concretely smartwatches and trackers. These technologies advance the 'quantified-self' approach and lead to more active and engaged health consumers and patients.[89] Comparatively, the Illumina/Grail case taps into the promise of genetic technologies and data analytics for facilitating personalised medicine.[90] As such, both cases showcase the potential of AI and digital technologies in increasingly data-driven healthcare and the promise of increased accessibility and quality of care while decreasing costs.[91]

### 3.3.1 Google/Fitbit Merger: An Approval, with Conditions

In December 2020, the Commission approved Google's acquisition of Fitbit with conditions.[92] Google is a provider of online search and advertising services, whereas Fitbit is a producer of connected wearable fitness devices and associated software,[93] including algorithms able to monitor users' wellness and detect different health conditions, like irregular heart rhythms.[94] According to Commissioner Vestager, the approval was based

88 Lynskey, *supra* n. 41, 217; see also Borgogno and Zangrandi, *supra* n. 73, 64.

89 Sharon, T. (2017). Self-tracking for health and the quantified self: Re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. *Philosophy & Technology*, *30*(1), 93-121.

90 See further Prainsack, B. (2020). Data mining in systems medicine and the project of solidarity: the interface of genomics and society revisited. In Mahr, D. & von Arx M. (eds.), *De-Sequencing. Health, Technology and Society* (Palgrave Macmillan Singapore 2020).

91 See further Terry, N. P. (2017). Appification, AI, and healthcare's new iron triangle. *Journal of Health Care Law & Policy*, *20*, 117-182.

92 Case M.9660 – *Google/Fitbit*, 14 August 2020. Note that the merger was also subject to procedures in other jurisdictions, including in Japan, South Africa, Australia, and the US; here, we only focus on the European Commission decision.

93 *Google LLC – Company Profile and News*, Bloomberg (2024), <https://www.bloomberg.com/profile/company/8888000D:US>.

94 *Fitbit's new heart monitoring algorithm reveals Google's plan to dethrone the Apple Watch*, The Next Web

on the commitments offered by Google to 'ensure that the market for wearables and the nascent digital health space will remain open and competitive'.[95] In its decision, the Commission assessed several concerns regarding competition. Here, we focus on those related to data, specifically Google having access to and control over Fitbit users' data and the technology for data collection and access.

It is posited that, in the Google/Fitbit merger, data collection played a key role due to both parties' business strategies.[96] This can be said in light of Google's overall mission to 'organize the world's information and make it universally accessible and useful' and Fitbit's claim that it is 'powered by one of the world's largest databases of activity, exercise and sleep data'.[97] In this light, one of the theories of harm focused on horizontal effects, expressing the Commission's concern that Google would combine user data from Fitbit – including step count, heart rate, sleep quality, calories burned, birth date, gender, height and weight data – with the vast array of data and capabilities Google already has, further improving Google's targeted advertising and strengthening its dominant position in online ad services and certain other data-based supply markets.[98] The Commission first stressed that there are regulatory limitations to prevent the illegal combining of datasets, including the EU's data protection and privacy rules. However, these 'do not eliminate the risks that the Parties' control on such data could render the expansion or entry by rival firms more difficult if not impossible'.[99] The Commission noted that Fitbit's data could be an important asset, for instance for profiling purposes.[100] It also noted that the data combination could improve Google's ability to personalise ads and that rivals without access to such data could face significant costs to compete.[101] In this light, the transaction could raise barriers to entry and expansion,[102] leading the Commission to conclude that

(12 April 2022), <https://thenextweb.com/news/fitbit-new-heart-monitoring-afib-algorithm-google-apple-smartwatch-analysis>.

95 *Mergers: Commission clears acquisition of Fitbit by Google, subject to conditions*, European Commission (17 December 2020), <https://ec.europa.eu/commission/presscorner/detail/en/ip_20_2484>.

96 Even though Google insisted that the deal was 'about devices, not data'. See Porter, J. & Statt, N. (2021) *Google completes purchase of Fitbit*, The Verge (14 January 2021), <https://www.theverge.com/2021/1/14/22188428/google-fitbit-acquisition-completed-approved>.

97 Salzberger and others, *supra* n. 47, 973.

98 Case M.9660 – *Google/Fitbit*, paras 399-402, 413.

99 Case M.9660 – *Google/Fitbit*, paras 403-13.

100 Case M.9660 – *Google/Fitbit*, paras 427-32.

101 Case M.9660 – *Google/Fitbit*, paras 445-53.

102 With regards to data combination and improved user profiling, the Commission also referred to responses of stakeholders who noted that the transaction would not only affect competition by actual or potential rivals, but also consumers more directly. This is because competition on 'digital markets takes place along various price and non-price parameters', including quality, innovation, and privacy, especially in light of users mostly paying with their data for 'free' services. Moreover, the merger would reduce the pressure for Google to compete on these non-price aspects, since it 'would further entrench Google's dominance'. Case M.9660 – *Google/Fitbit*, paras 451-2.

the merger was likely to negatively affect competition in markets for online advertising.[103] Furthermore, the Commission focused on the role of Fitbit data as input for various digital health apps, in particular on the possibility for an input foreclosure (vertical effects). This is in light of businesses in the digital health field that derive Fitbit user data from its APIs.[104] The Commission suggested that digital health is a nascent, fragmented sector, which is expected to grow. Restricting access to Fitbit API may affect digital health players' success and their contribution to innovation and diversification of the digital healthcare sector, thus leading to the finding that the concerns over 'a significant detrimental effect on competition in the digital healthcare sector if the merged entity would restrict access to Fitbit's Web API' cannot be discarded.[105]

To remedy the competitive concerns, the following behavioural commitments were envisioned. With regards to data combination issues, Google agreed to silo Fitbit data and to not use it for online advertising purposes. The commitments also include an obligation that users need to be presented with a choice to grant or deny the use of data on their body functions, activities and condition for other Google services.[106] Concerning access to input, Google committed to continuing to make Fitbit's API available without charging for access, subject to relevant terms of service and certain privacy and security requirements.[107] These commitments are in place for ten years and their implementation will be monitored by an appointed trustee.[108]

While the imposition of a data silo is quite far-reaching from a competition perspective, the possible privacy impact of the merger, which was not explicitly assessed by the Commission, raised extensive attention. Fears were raised about the sensitivity of the health data that would fall into Google's hands. Concerns were also expressed about the future impact of the merger on the overall development of healthcare and insurance, beyond the more narrow markets of advertising and digital health that the Commission assessed.[109] Interestingly, Google committed to putting a data protection system in place

103 Case M.9660 – *Google/Fitbit*, paras 454-5. Conversely, the Commission found that Fitbit's data is less relevant in the market for general search and that the merger would not negatively impact competition thereon. A similar conclusion was reached concerning the market for digital healthcare, where several well-established alternatives are present. Case M.9660 – *Google/Fitbit*, paras 472-4 and 488-96.

104 Case M.9660 – *Google/Fitbit*, paras 503-4.

105 Case M.9660 – *Google/Fitbit*, paras 526-31.

106 Case M.9660 – *Google/Fitbit*, paras 944, 964-73.

107 Case M.9660 – *Google/Fitbit*, paras 945-51, 974-84.

108 See also European Commission, *supra* n. 96. For further discussion, see also Simon Vande Walle, The European Commission's Approval of Google/Fitbit – A Case Note and Comment, Concurrences Competition Law Review Nr. 3-2021 (2021), <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3893079>.

109 See, e.g., Bria, F. and others (2020), *Europe must not rush Google-Fitbit deal*, Politico (22 July 2020), <https://www.politico.eu/article/europe-must-not-rush-google-fitbit-deal-data-privacy/>.

to implement the data silo and to ensure that Google and Fitbit's datasets remain separate.[110] This illustrates how an economic concern relating to data concentration can be remedied by using mechanisms from legal domains protecting non-economic interests, like privacy, with regard to personalised online advertising. However, the question remains as to what Google's plans are with using Fitbit data in its non-advertising businesses.[111] These business endeavours among other things include developing AI for healthcare,[112] where non-economic data-related harms could negatively affect health policy objectives other than economic aspects.[113]

### 3.3.2 Illumina/Grail Merger: A Prohibition

In September 2022, after an in-depth investigation, the Commission prohibited[114] Illumina's implemented acquisition of Grail due to concerns that the merger would have 'stifled innovation, and reduced choice in the emerging market for blood-based early

110 Case M.9660 – *Google/Fitbit*, Commitments to the European Commission, 4 November 2020, Section A.1 sub d: 'To the extent that a Google Service accesses Measured Body Data or Health and Fitness Activity Location Data, Google will apply a Data Protection System to ensure Data Separation of the accessed data'. Note, however, that this only concerns Google's commitments regarding ads. In this regard, the Australian competition authority (ACCC) chair further expressed concerns over how to monitor and enforce the behavioural commitments to not use Fitbit data for advertising. As pointed out, the concern seems apt in the context of big tech firms 'reaping the benefits of consolidating two lucrative datasets' despite previous claims – as also appeared was the case in the Facebook/WhatsApp merger or in the Google/DoubleClick cases. See Panichi, J. (2021) *Australia Goes it alone*, Inside Story (9 April 2021), <https://insidestory.org.au/australia-goes-it-alone/>.

111 Kemp, K. (2020) *Every step you take: why Google's plan to buy Fitbit has the ACCC's pulse racing*, The Conversation (23 June 2020), <https://theconversation.com/every-step-you-take-why-googles-plan-to-buy-fitbit-has-the-acccs-pulse-racing-141052>.

112 Landi, H. (2023) *Google ramps up AI tech for health tools, app development*, FierceHealthcare (14 March 2023), <https://www.fiercehealthcare.com/health-tech/google-launches-open-health-stack-app-developers-unveils-new-ai-partnerships>; Kingson, J. A. (2023). *Google flexes its health care AI muscle*, Axios (15 March 2023), <https://www.axios.com/2023/03/15/google-ai-chatgpt-healthcare-youtube-chatbot-health>; Nieva, R. *Google Will Let Healthcare Organizations Use Its AI To Analyze And Store X-Rays*, Forbes (4 October 2022), <https://www.forbes.com/sites/richardnieva/2022/10/04/google-cloud-medical-imaging-x-rays/>.

113 In addition to undermining the right to privacy, human rights advocates further warned that Google's surveillance-based business model could hamper rights to health 'where people may suffer unequal treatment based on predictions about their health, and as such must be taken into account in the context of health and fitness data'. Amnesty, *supra* n. 51. See also Lomas, N. (2020) *Google gobbling Fitbit is a major privacy risk, warns EU data protection advisor*, TechCrunch (20 February 2020), <https://techcrunch.com/2020/02/20/google-gobbling-fitbit-is-a-major-privacy-risk-warns-eu-data-protection-advisor>.

114 The full text of the decision with the number M.10188 is, at the time of writing, not yet available in the public register of competition cases; therefore, we were only able to review the Commission's – quite detailed – press release. See *Mergers: Commission prohibits acquisition of GRAIL by Illumina*, European Commission (6 September 2022), <https://ec.europa.eu/commission/presscorner/detail/en/ip_22_5364>.

cancer detection tests'.[115] Illumina develops, produces and markets genetic sequencing systems and other technologies that increasingly rely on AI to process large amounts of data and detect patterns in DNA.[116] Grail, which was founded by Illumina in 2015 and later spun off, is focused on early cancer detection technologies that use AI to analyse blood samples.[117] As noted, Illumina is currently the only supplier of the next-generation sequencing (NGS) technology required to develop these tests. Market players expressed concerns that Illumina was trying to gain control over the emerging early cancer detection testing market.[118] The investigation established that the merger would indeed grant the firm an incentive and the ability to foreclose or otherwise disadvantage innovation competition in this domain. This would be the case even if the benefits would be realised at a later stage,[119] as Grail's rivals are dependent on Illumina's NGS technology to develop and market their own tests, hence preventing the development of such blood tests by others.[120] By, for instance, refusing to supply its NGS technology to rivals, price increases, degradation of quality or delaying supplies with no credible alternatives and significant barriers to entry, Illumina could negatively affect the 'innovation race' to develop and commercialise early cancer detection tests in the EU that are predicted to become highly lucrative and poised to 'revolutionise our fight against cancer and help to save millions of lives'.[121] These concerns were deemed not sufficiently addressed by the remedies proposed by Illumina, leading to the Commission prohibiting the transaction. As the merger had already been implemented, the Commission proposed measures in December 2022 for Illumina to unwind the acquisition of Grail.[122]

---

115 European Commission, *supra* n. 114.

116 *AI drives data insights for doctors and their patients*, Illumina (13 June 2022), <https://www.illumina.com/company/news-center/feature-articles/ai-drives-data-insights-for-doctors-and-their-patients.html>.

117 Lipsky, T. (2022). *The Neo-Brandeisian Approach to Vertical Mergers—A Zipline to Oblivion?*, CPI Antitrust Chronicle (November 2022), <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4281090>. Smith, W. (2020). *Illumina acquires ML-powered cancer detection firm Grail*, AIMagazine (22 September 2020), <https://aimagazine.com/ ai-applications/illumina-acquires-ml-powered-cancer-detection-firm-grail>.

118 European Commission, *supra* n. 114.

119 Falce, V. and Faraone, N.M.F (2023) Digital Ecosystems in the Wake of a Legislative/Regulatory Turmoil: A First (Tentative) Antitrust Assessment of the Italian (and European) Experience in the AGCM Case Law. *World Competition. Law and Economics Review*, *46*(1), 37-64.

120 European Commission, *supra* n. 114.

121 European Commission, *supra* n. 114; see also Falce and Faraone, *supra* n. 119, 58-9.

122 Chee, F. Y. (2022) *EU charge sheet tells Illumina to swiftly unwind Grail deal*, Reuters (5 December 2022), <https://www.reuters.com/markets/deals/illumina-gets-eu-antitrust-charge-sheet-unwind-grail-acquisition-swiftly-2022-12-05/>; see further Chee, F. Y. *Illumina challenges EU order to keep Grail separate*, Reuters (10 January 2023), <https://www.reuters.com/markets/deals/illumina-challenges-eu-order-keep-grail-separate-2023-01-10/>. At the end of 2023, the Commission ordered Illumina to unwind the acquisition of Grail by adopting restorative measures under the EU Merger Regulation; concretely, it ordered Illumina to adopt divestment and certain transitional measures. *Commission orders Illumina to unwind its completed acquisition of GRAIL*, European Commission (12

Illumina's acquisition of Grail and the merger control procedure that followed was interesting from several perspectives.[123] The focus here is on its substantive approach, specifically the dimensions linked to data and AI. While access to and the accumulation of data was arguably one of the matters at the forefront of the Google/Fitbit merger, the Illumina/Grail case seemingly did not tackle data-related issues directly. We nonetheless consider the case relevant for the present discussion since the competitive concerns in the decision revolve around AI and data-driven technologies for cancer detection. In oncology care and research, NGS technologies that generate and analyse large datasets, could significantly contribute to finding novel biomarkers that could help improve cancer detection, diagnosis and treatments.[124] In this light, the merger decision that attempted to ensure and preserve data-driven innovation contributed to the realisation of an important policy objective in healthcare by treating it in economic terms. Specifically, beating cancer is 'a main priority in the area of health

October 2023), <https://ec.europa.eu/commission/presscorner/detail/en/ip_23_4872>. In the US, the FTC similarly ordered Illumina in April 2023 to divest Grail in order to 'protect competition in life-saving technology market'. *FTC Orders Illumina to Divest Cancer Detection Test Maker GRAIL to Protect Competition in Life-Saving Technology Market*, FTC (3 April 2023), <https://www.ftc.gov/news-events/news/press-releases/2023/04/ftc-orders-illumina-divest-cancer-detection-test-maker-grail-protect-competition-life-saving>.

123 See further *Mergers: Commission alleges Illumina and GRAIL breached EU merger rules by early implementation of their acquisition*, European Commission (19 July 2022), <https://ec.europa.eu/commission/presscorner/detail/en/ip_22_4604>; *Mergers: Commission adopts interim measures to prevent harm to competition following Illumina's early acquisition of GRAIL*, European Commission (29 October 2021), <https://ec.europa.eu/commission/presscorner/detail/en/ip_21_5661>. In particular, the case was important for the EU merger control doctrine also because it was the first time the Commission reviewed and blocked a transaction that did not meet the EU Merger Regulation and was not notified under Member States' local merger control regimes. The Commission explicitly invited referral requests by Member States under Article 22 of the EU Merger Regulation; subsequently, the referral was made by France, joined by Belgium, Greece, Iceland, the Netherlands and Norway. The Commission deemed that the referral was appropriate it would affect trade in the single market and because Grail's competitive significance is not reflected in its turnover. *Illumina/GRAIL: EC Blocks Transaction Below EU and Referring Member State Merger Control Thresholds for the First Time*, Cleary Gottlieb (15 September 2022), <https://www.clearygottlieb.com/news-and-insights/publication-listing/illumina-grail-ec-blocks-transaction-below-eu-and-referring-member-state-merger-control-thresholds-for-the-first-time>; Falce and Faraone, *supra* n. 119, 59; European Commission, *supra* n. 114; *The General Court upholds the decisions of the Commission accepting a referral request from France, as joined by other Member States, asking it to assess the proposed acquisition of Grail by Illumina*, Court of Justice of the European Union (13 July 2022), <https://curia.europa.eu/jcms/upload/docs/application/pdf/2022-07/cp220123en.pdf>. Please note that while this chapter was in press, on 3 September 2024, the Court of Justice annulled the Commission's decision(s) to examine the acquisition and set aside the ruling of the General Court in the Illumina/Grail merger. We still consider the approach to the substantive assessment worthy of discussion.

124 Dlamini, Z., Francies, F. Z., Hull, R., & Marima, R. (2020). Artificial intelligence (AI) and big data in cancer and precision oncology. *Computational and structural biotechnology journal*, *18*, 2300-2311.

of the von der Leyen Commission'[125] and the EU announced a strategy 'to leave no stone unturned to take action against cancer'.[126] The strategy stresses the importance of health data and new technologies for developing personalised medicine, among other things.[127] Improving the early detection of cancer, including promoting novel approaches, is among the key areas[128] since it can lead to better health outcomes.[129] From this perspective, it can be observed how competition enforcement reinforces or acts in harmony with the objectives of health policy – the development of innovative technologies in this case – by prohibiting market conduct deemed to violate these objectives. This seems consistent with competition enforcement in pharmaceutical markets, where EU merger control has been able to contribute 'to improving European patients' access to affordable and innovative essential medicines'[130] as policy goals with significant non-economic implications.

As such, even though this aspect remains implicit in its decision, the Commission has arguably contributed to health data's value and utility for innovation and to achieving healthcare objectives beyond efficiency by applying its regular merger assessment that focuses on addressing competition concerns. The Illumina/Grail case therefore shows that where economic and non-economic interests overlap, competition remedies that are imposed to address economic harm can also have a spill-over effect and remedy possible non-economic harm. The difference with the Google/Fitbit case is that the competition concerns in the latter were not considered problematic enough to justify a further-reaching remedy than a data silo or even a prohibition, as in Illumina/Grail. The question that these cases raise is to what extent merger control can be used to impose remedies that go beyond what is necessary to address economic concerns and can remedy non-economic concerns, such as considerations of privacy and equitable (access to) healthcare, as discussed in the next section.

125 *Europe's Beating Cancer Plan: A new EU approach to prevention, treatment and care*, European Commission (3 February 2021), <https://ec.europa.eu/commission/presscorner/detail/en/ip_21_342>.

126 *Europe's Beating Cancer Plan. Communication from the Commission to the European Parliament and the Council (2022), <https://health.ec.europa.eu/system/files/2022-02/eu_cancer-plan_en_0.pdf>*, 4.

127 Ibid., 6.

128 Ibid, 14-5.

129 Smith, *supra* n. 117.

130 *Competition: Commission report finds active competition enforcement contributes to affordable and innovative medicines*, European Commission (28 January 2019), <https://ec.europa.eu/commission/presscorner/api/files/document/print/en/ip_19_741/IP_19_741_EN.pdf>.

## 4. An Analysis: The Scope for Non-Economic Considerations in Merger Control

Despite their outcomes, the two decisions have not explicitly engaged with non-economic data-related harms, nor with the relationship between competition and healthcare objectives.[131] Recently, however, the calls to include 'matters of public interest' in examining the impact of concentrations on the internal market have become louder.[132] In this final part, we explore the possibility of EU merger control to ensure that technology and digital transformation truly work for everyone[133] and discuss the motivation and potential avenues for integrating non-economic aspects into merger analyses.

### 4.1 Motivations for Integrating Non-Economic Considerations into EU Merger Control[134]

'[Data] sharing is not a value-neutral practice, but rather represents a broad spectrum of ethical, political and social goals that various actors are seeking to achieve'.[135] As such, we argue that there is room for merger control to address non-economic considerations resulting from data-driven healthcare mergers more proactively. While critics may argue that competition authorities should only include economic interests in their analysis, we posit that the current reality shows that non-economic impacts can no longer be fully separated from economic considerations. Accordingly, we explore how and to what extent non-economic concerns can be considered in competition assessments of data-driven mergers in healthcare.[136] These can, as elaborated earlier, affect the functioning of the market, but also have broader harmful effects that arguably cannot be easily decoupled from market activity. Accordingly, it is logical to look into data sharing objec-

131 See also Hancher and Sauter, *supra* n. 66, 251.

132 See, e.g., European Parliament, *Draft report on competition policy – annual report 2022*, 2022/2060(INI)

133 *Europe's digital future*, European Commission (2024), <https://competition-policy.ec.europa.eu/about/europes-digital-future_en>.

134 The discussion in section 4.1 is partly based on the forthcoming PhD manuscript of Tjaša Petročnik.

135 Riso and others, *supra* n. 7, 12.

136 See, e.g., the opinion of the EDPB that suggests that longer-term implications for the protection of economic, data protection, and consumer rights need to be assessed whenever a significant merger is proposed, especially in the technology sectors, as increased market concentration has in digital markets has the potential to threaten the level of data protection and freedom enjoyed by consumers of digital services. *Statement of the EDPB on the data protection impacts of economic concentration*, European Data Protection Board (27 August 2018), <https://www.edpb.europa.eu/our-work-tools/our-documents/other-guidance/statement-edpb-data-protection-impacts-economic_en>.

tives and purposes and how they contribute to efficiency and equity goals in healthcare, thus basing the analysis on qualitative values rather than only quantitative metrics.[137]

While the current interpretation of merger control by the Commission mainly focuses on the maximisation of economic efficiency, EU competition law has always encompassed considerations wider than that. Competition rules have been concerned with the proper functioning of the EU internal market and are intended to 'prevent competition from being distorted to the detriment of the public interest, individual undertakings and consumers, thereby ensuring the well-being of the European Union'.[138] In this sense, a broader understanding of EU competition law and, in this case, merger control would establish a tighter link between the market and its societal context,[139] counter the concentration of power that might harm the democratic process (to which sufficiently 'efficient' mergers might lead),[140] and promote values like market freedom, social justice and solidarity.[141] An additional justifying argument is based on the EU's constitutional framework that foresees EU competition law being interpreted in light of various Treaty objectives and principles.[142] The Treaty namely includes various 'integration clauses' that mandate that, in designing and implementing its policies, the EU considers particular values and principles,[143] and requires consistency in all its policies and activities.[144] This indicates that EU competition law simply cannot be separate from certain (other) 'public interest' considerations recognised in the Treaty[145] that are less economic in

137 Eben, M. (2018). Angela Daly, Private Power, Online Information Flows and EU Law: Mind the Gap. *International Data Privacy Law, 8*(3), 284–287.

138 Case C-52/09 *TeliaSonera* ECLI:EU:C:2011:83, para 22.

139 Majcher, K., & Robertson, V. H. (2022). Protecting Personal Data and the Environment: Doctrinal Challenges for EU Competition Law in our Day and Age. *European law review*, *47*(5), 622-646, referring to the Ordoliberal ideas.

140 Eben, *supra* n. 137, 285.

141 Wörsdörfer, M. (2020). Ordoliberalism 2.0: Towards a new regulatory policy for the digital age. *Philosophy of Management*, *19*, 191-215.

142 Witt, A. C. (2012). Public Policy Goals Under EU Competition Law—Now is the Time to Set the House in Order. *European Competition Journal*, *8*(3), 443-471; Majcher and Robertson, *supra* n. 139.

143 See Articles 9-13 TFEU; in relation to healthcare specifically, Article 9 is relevant: 'In defining and implementing its policies and activities, the Union shall take into account requirements linked to the promotion of a high level of employment, the guarantee of adequate social protection, the fight against social exclusion, and a high level of education, training and protection of human health'.

144 Article 7 TFEU, which states that '[t]he Union shall ensure consistency between its policies and activities, taking all of its objectives into account and in accordance with the principle of conferral of powers'. See also Article 13 TEU.

145 Iacovides, M. C., & Vrettos, C. (2022). Falling through the cracks no more? Article 102 TFEU and sustainability: the relation between dominance, environmental degradation, and social injustice. *Journal of Antitrust Enforcement*, *10*(1), 32-62; see also Ezrachi, A. (2017). Sponge. *Journal of Antitrust Enforcement*, 5(1), 49-75. In this light, Lianos, Minssen, and Kollmar point to the regulatory osmosis, namely the absorption of (other) regulatory aims into competition law enforcement. Lianos, I.,

nature. When it comes to healthcare, it nonetheless needs to be acknowledged that, according to the Treaty, the organisation of healthcare systems and policy is in the hands of Member States. Still, there could be a possibility for the EU to act upon the broader societal values in this domain through its economic powers, as has happened before.[146] This provides a further justification for examining the role and scope for EU merger control in relation to upholding both the economic and non-economic objectives associated with healthcare domains when addressing data-driven business strategies.

### 4.2 Two Scenarios: Lessons from the Analysed Mergers

The competition law framework, and merger control specifically, does not currently consider the non-economic impact of acquisitions or data aggregation agreements on individuals, much less as regards its social impact.[147] Nevertheless, the Google/Fitbit and Illumina/Grail cases show that economic and non-economic interests overlap and that merger remedies may sometimes also (partly) address non-economic harms resulting from mergers. Two scenarios can be distinguished: (1) the merger does not significantly impede effective competition, and (2) the merger does significantly impede effective competition.[148]

In the first scenario where competition concerns are absent, the Commission has to declare the merger compatible with the internal market because its competences to impose remedies or block mergers stemming from the EU Merger Regulation do not apply if there is no significant impediment to effective competition. In these cases, however, Article 21(4) of the EU Merger Regulation provides that a merger may be sent to a Member State for further assessment on non-competition grounds. Namely, Member States might want to protect 'legitimate interests other than those taken into consideration by the Regulation and compatible with the general principles and other provisions of EU law', including public security, plurality of the media and prudential rules.[149] In such cases, the consideration of purely economic considerations can be

Minssen, T., & Kollmar, C. (2022). Tackling grand challenges with competition law: lessons from the pandemic. In W. Sauter, M. Canoy, and J. Mulder (Eds.). *EU Competition Law and Pharmaceuticals*. Edward Elgar Publishing.

146 Majcher and Robertson, *supra* n. 139; see also Brooks, E., de Ruijter, A., & Greer, S. L. (2020). COVID-19 and European Union health policy: From crisis to collective action. In Vanhercke, B., Spasova, S. & Fronteddu, B. (eds.), *Social Policy in the European Union. State of Play 2020: Facing the pandemic – Twenty-first annual report* (European Trade Union Institute 2021).

147 Lynskey, *supra* n. 41, 218.

148 These scenarios stem from the discussion in Graef, I. (2016). *EU Competition Law, Data Protection and Online Platforms: Data as Essential Facility: Data as Essential Facility*. Kluwer Law International BV.

149 Article 21(4) EU Merger Regulation.

excluded to preserve a particular value, and mergers could be prohibited on public interest grounds.[150]

This analysis takes place outside the framework of EU merger reviews and is based on national law. As a result, the initiative for performing such a review lies with the Member States and their respective competent authorities. This also signals the limits of EU merger control and indicates the shared responsibilities different regulators have in protecting against the harms resulting from data-driven mergers. As argued elsewhere, national data protection authorities can use Article 21(4) of the EU Merger Regulation to conduct a data protection assessment of data-driven mergers and align their assessments with the Commission's regular and more economic merger assessments.[151] While there is scope for the Commission to be more proactive in addressing non-economic interests when competition concerns are present, as discussed below, data protection and healthcare authorities also have a role to play if the Commission is not competent to act because no competition concerns are identified. The room that the EU Merger Regulation offers Member States to do so is thus far unused in the context of data-driven mergers.[152] This may be explained by the cautious approach of Member States not to interfere with the Commission's assessment of such mergers.

In the second scenario, competition concerns are present and the Commission is competent to impose remedies or even block the merger. If a proposed merger is found to significantly impede effective competition, the merging parties typically offer remedies to the Commission to address the identified competition concerns and to ensure the merger can be made compatible with the internal market. Because the merging parties are dependent on the discretion of the Commission to prevent the merger from being blocked, we argue that the Commission has the scope to require remedies that go beyond ending the identified competition concerns. The remedies must of course be proportionate to the competition concerns. However, the fact that non-economic interests are increasingly intertwined with economic considerations provides a scope for

150 Lynskey, *supra* n. 41, 218-9. As a point of curiosity, South African merger control regime allows prohibitions of mergers based on public interest grounds; in this, Burger King is considered a precedent decision that prohibited the transaction on the basis that it will have a substantial negative effect on 'the promotion of greater spread of ownership, in particular to increase the levels of ownership by historically disadvantaged persons'. *Precedent-Setting Decision: Burger King Acquisition Prohibited Purely on Public Interest Grounds*, CPI Blogs (1 July 2021), <https://www.pymnts.com/cpi-posts/precedent-setting-decision-burger-king-acquisition-prohibited-purely-on-public-interest-grounds/>.

151 Graef, I., Clifford, D., & Valcke, P. (2018). Fairness and enforcement: bridging competition, data protection, and consumer law. *International Data Privacy Law*, *8*(3), 200-223.

152 Examples of cases where Article 21(4) of the EU Merger Regulation has been invoked is by the UK for conducting a media plurality review of mergers on the basis of national law. See for instance *Mergers: Commission clears 21st Century Fox's proposed acquisition of Sky under EU merger rules**, European Commission (7 April 2017), <https://ec.europa.eu/commission/presscorner/detail/en/IP_17_902>.

– or even demands – a more proactive approach. This is arguably especially the case when considerations of fundamental rights are at play.

Privacy and health protection interests are protected as fundamental rights in the EU legal order. Articles 7 and 8 of the EU Charter of Fundamental Rights lay down the rights to privacy and data protection. With regard to healthcare, Article 35 not only establishes a right of access to preventive healthcare and a right to benefit from medical treatment but also requires a high level of human health protection to be ensured 'in the definition and implementation of all Union policies and activities'. More generally, Article 51(1) of the Charter requires Union institutions to 'respect the rights [laid down in the Charter], observe the principles and promote the application thereof in accordance with their respective powers'. Following these provisions, one can question whether the Commission may be under a duty to proactively promote the rights and interests protected in the Charter when exercising its competences under the Treaties and EU legislation.[153]

Article 51(2) of the Charter makes clear that its provisions do not establish any new powers or tasks for the EU, or modify powers and tasks defined by the Treaties. Nonetheless, when conducting merger assessments based on the EU Merger Regulation, the Charter may be interpreted in a way that expects the Commission to guarantee the effectiveness of fundamental rights, like those involving privacy, data protection, and health protection, by adopting measures that actively promote non-economic interests relating to these fundamental rights and go beyond what is necessary to protect against economic harm. In practice, this would mean that the Commission's merger assessments in cases where competition concerns are present should also proactively consider related impacts on non-economic interests protected in the Charter. It is uncertain how far this responsibility and accountability stretches. However, it does indicate that there is leeway and perhaps even a duty for the Commission to pay more attention to non-economic interests in merger cases than it currently does. The Commission could do such assessment itself or it could collaborate with the relevant authorities, or even outsource part of the assessment, where it feels it lacks the necessary expertise.

The more proactive use of Article 21(4) of the EU Merger Regulation in both of the scenarios considered here could be an initial starting point for aligning a more economic merger assessment with an assessment of non-economic interests. This provision, however, still presumes a separate assessment of each. The current reality in healthcare and other data-driven industries where economic and non-economic interests are becoming increasingly intertwined demands more integrative approaches that were not yet foreseen when the EU Merger Regulation was adopted. However, because the possibilities offered by the existing merger framework are not used by the different

153 In the context of the right to privacy and data protection, see the discussion in Graef, *supra* n. 148, 343-5.

responsible authorities yet, a first priority would be to encourage stronger and more effective cooperation and coordination. By doing so, the Commission's existing economic analysis of mergers can be complemented by reviews of the effects of mergers on privacy or health protection by the respective authorities. This would make any non-economic impacts of data-driven mergers in the health sector more visible. Depending on their competences, the respective authorities may also be able to take relevant measures under national law to monitor or address the adverse effects of a merger, regardless of whether it raises competition concerns.

## 5. Conclusion

In 2022, the Chairwoman of the US FTC Lina Khan wrote that the goal of antitrust authorities, at least in the US, is 'to prevent illegal mergers, not to make the world a better place'.[154] Despite the differences between merger control in the US and EU, this statement seems to hold true for the EU as well, considering the focus of the Commission's analysis in the merger cases discussed in this chapter. However, data-driven mergers in healthcare take place in a context where the overarching values of universality, access to good quality care, equity and solidarity[155] coexist with the calls for efficiency, cost control, affordable treatments, patient choice and innovation.[156] Data technologies and AI are commonly considered to be disruptive to these goals,[157] in both positive and negative senses. We have thus considered it necessary to examine whether merger control in the EU, due to its political and constitutional tenets, could accommodate merger assessments that consider economic (efficiency) issues and the impediment of effective competition, together with non-economic (equity) concerns like privacy violations or algorithmic bias. As suggested, this approach can raise questions about increased complexity and legal uncertainty.[158] However, it could importantly contribute to greater consistency in policies as well, especially when market activity can affect health.

We have shown that, while there is scope for the Commission to be more proactive in imposing merger remedies that could also protect non-economic interests, there are limits to the competences of the Commission under the EU Merger Regulation and that there is a need for data protection and healthcare authorities to be involved as well. The

154 Khan, L. (2022). *ESG Won't Stop the FTC*, WSJ (21 December 2022), <https://www.wsj.com/articles/esg-wont-stop-the-ftc-competition-merger-lina-khan-social-economic-promises-court-11671637135>.

155 Council Conclusions on Common values and principles in European Union Health Systems, OJ C 146, 22 June 2006.

156 European Commission, *supra* n. 130.

157 See further Terry, *supra* n. 91.

158 See, e.g., Deutscher, *supra* n. 78.

existing framework already offers opportunities to do so, but these have not been explored in practice yet. Because of the increasing overlap and interaction between economic and non-economic interests in data-driven industries such as healthcare, more experimentation and coordination by the different regulatory authorities will be vital to ensure that the benefits of AI and data are maximised and that possible harms are adequately identified and addressed.

CHAPTER

# X

# Law-Making, Knowledge-Making, World-Making: Reading the EU AI Act Through an Epistemic In/Justice Lens

**Aviva de Groot &**

**Siddharth Peter de Souza**[1]

https://doi.org/10.26116/m72t-4v82

1 Postdoc Tilburg Institute of Law, Technology, and the Society (TILT), Tilburg Law School (TLS), Tilburg University. Research Associate at ACEPS — The African Centre for Epistemology and Philosophy of Science" to our postdoc positions

## 1. Introduction

In domains as diverse as law enforcement, medicine, news services, and public policy, a major pitfall of the use of artificial intelligence (AI) is its propensity to reproduce racist, sexist, casteist and other harmful outcomes that already permeate through societal structures.[2] Studies of the history and historical actors of the computational fields reveal the longevity of such problems, and testify to how reductive and even oppressive moral notions about humanity and the world were of defining influence in the fields' development.[3] Katz for example argues how computer science has flexibly served 'whiteness's' imperial projects with 'epistemic forgeries', depicting racial visions of the world as authentic and representative.[4] This flexibility, he argues, shows in how the technically undefined umbrella term 'AI' can effectively rebrand into for example 'AI for Good' in ways that obfuscate these problematic ideological underpinnings, leaving them in a place where they continue to do harm.[5]

Studies highlight different dimensions to the computational fields' problematic character. Some focus on the type of knowledge that methods like machine learning allow to create with their input, and how systems using them cannot simply be repurposed towards more beneficial ends;[6] others focus on the problematic hegemony of the industry, its dominance over computational infrastructures, its geographical concentration in just a few parts of the world and thereby the exclusion of experts, voices, and contexts from the majority worlds.[7] Attention is also called to 'ignorance making' by the field in how it trains its new recruits. For example, Rankin et al. write how Computer

2 Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press. Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. John Wiley & Sons. Bokil, A. and others (2021), 'Settled Habits, New Tricks: Casteist Policing Meets Big Tech in India' (Longreads) <https://longreads.tni.org/stateofpower/settled-habits-new-tricks-casteist-policing-meets-big-tech-in-india> accessed 30 May 2021.

3 Lepore, J. (2020). *If then: How the simulmatics corporation invented the future*. Liveright Publishing. Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. MIT Press. Whittaker, M. (2023). Origin stories: Plantations, computers, and industrial control. *Logic (s)*, *19*.

4 Katz, Y. (2020). *Artificial whiteness: Politics and ideology in artificial intelligence*. Columbia University Press.

5 Some readers may find these opening statements unduly negative. We ask to consider how arguments, even critical ones, on AI's legal treatment typically start with characterizations that boil down to 'good in potential, bad in cases'. As we will argue in this chapter, a more honest narrative starts from AI trouble. Furthermore, the AI field is neither short of defenders nor funding for them and for business. Those that are harmed by AI lack both, and we see it as a legal duty to amplify their voices.

6 For example, Malik, M. M. (2020). A hierarchy of limitations in machine learning. *arXiv preprint arXiv:2002.05193*.

7 Amrute, S., Singh, R., & Guzmán, R. L. (2022). A primer on AI in/from the Majority World: An Empirical Site and a Standpoint. *Available at SSRN 4199467*.

Science education has not acknowledged deeply rooted discriminatory practices, including its gatekeeping of who can produce acceptable knowledge, controlling what Collins called 'epistemic power' and restricting it to dominant and privileged interpretive communities. A false narrative has thus emerged of Computer Science as a 'colourblind and meritocratic discipline'.[8]

In times of hype cycles about 'AI's positive and negative potential'[9] and in light of big tech's spins ideas about what that should mean for regulation,[10] it is all the more important to acknowledge how the harms of AI already manifest, and are not as dependent on either novelty or cutting-edge complexity as it is argued. Placing AI-based systems in their historical and actual contexts can help to ensure that they are crafted differently moving forward, rather than used to further the societal structural inequalities that they are a product of. This is demonstrated in cases such as the COMPAS recidivism risk algorithm in the United States[11] and that of the Dutch municipality of Rotterdam's welfare fraud prediction system.[12]

This chapter argues that the regulation of AI by the EU lawmaker provides a particular momentum for such critical engagement. As some in the digital rights field argue, it is not just necessary to legally protect against AI harms, but to do so in a way that includes those voices and standpoints that have not been represented in technology regulation so far.[13] Increasingly, a multidisciplinary research community interested in transforming AI into a more justice-oriented field that produces insights that can use-

8 Rankin, Y. A., Thomas, J. O., & Erete, S. (2021). Black women speak: Examining power, privilege, and identity in CS education. *ACM Transactions on Computing Education (TOCE)*, *21*(4), 1-31.

9 See for example, 'Statement on AI Risk: AI experts and public figures express their concern about AI risk' (Center for AI Safety, originally posted on May 30, 2023 <https://www.safe.ai/statement-on-ai-risk>); 'The Godfather of A.I.' Leaves Google and Warns of Danger Ahead, (Ny Times, May 1 2023 <https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html>)

10 'Big Tech's Heavy Hand Around the Globe' (*Human Rights Watch*, 8 September 2020) <https://www.hrw.org/news/2020/09/08/big-techs-heavy-hand-around-globe> accessed 13 March 2023. 'Google, Microsoft, OpenAI and startup form body to regulate AI development' (The Guardian, 26 July 2023 <https://www.theguardian.com/technology/2023/jul/26/google-microsoft-openai-anthropic-ai-frontier-model-forum>).

11 Larson et al, How We Analyzed the COMPAS Recidivism Algorithm (ProPublica May 2016 <https://www.propublica.org/Article/how-we-analyzed-the-compas-recidivism-algorithm?token=X-SO7CCiM7DoudJrFYQeZnvAitR3ZT0sj>)

12 Suspicion Machines: Unprecedented experiment on welfare surveillance algorithm reveals discrimination (Lighthouse Reports, https://www.lighthousereports.com/investigation/suspicion-machines/)

13 Nani Jansen Reventlow Rebuilding the master's house instead of repairing the cracks: why "diversity and inclusion" in the digital rights field is not enough (Digital Freedom Fund September 2029, <https://digitalfreedomfund.org/rebuilding-the-masters-house-instead-of-repairing-the-cracks-why-diversity-and-inclusion-in-the-digital-rights-field-is-not-enough/>); EDRi, Decolonising digital rights, <https://edri.org/what-we-do/decolonising-digital-rights/>.

fully inform regulation.[14] In this chapter, we argue that the AI Act provides an important opportunity for the law to improve the quality of its efforts, and for legal scholars to step up their own efforts while informing, guiding, and criticizing the ways in which AI's governance and regulation is envisaged in the work of the EU.

We also argue that this momentum comes with special characteristics that are tied to Europe's historically problematic role as norm-setter for both regulatory and knowledge-making practices, and its dominant position today in the field of technology regulation.[15] European colonial expansion was supported by the development of knowledge and governing systems that allowed it to systematize oppression, informed by scientific focuses and scientific practices that sustained these aims.[16] These effects are felt today in, among others, economic, bureaucratic, legal and scientific realities – including AI.

Importantly, the entanglement of these practices also emerge in how law defines and imagines the regulated world, the people in it, and its own role in it, including institutional requirements it imposes or the types of organizations it excludes. Lawmaking is in this way an epistemic endeavour in itself, traditionally entailing extensive debates on concepts, categories, rules on evidence, expertise and credibility, among other subjects. And it is here that we locate a point for engagement with the parallel of knowledge-making and law-making.

With this in mind, we are interested in studying the role of law in the governance of AI and the protection of people from harmful output of AI systems. If law has not put any meaningful constraints in place to limit AI-driven harms, as seems to be the case, can it be argued that law has been a facilitator, and if so, in what way(s)? We know that law, too, has flexibly served the interests of societies based on white supremacy and that it has been an instrument of exclusion, perpetuating violence through colonial practices, for example by establishing 'civilising missions', over centuries.[17] Like AI, law too has been able to respond to and to some extent absorb critical voices without losing its usefulness for those who wish to use it towards oppressive ends.[18] Can we pinpoint useful aspects of this parallel to work towards a more equitable governance of AI?

---

14 See for example <https://facctconference.org/>.

15 This statement does not mean to do disservice to good things that have come out of EU's regulatory engagement with data science, as expressed in a broad range of first-of-their-kind efforts to regulate against among other things the misuse of personal data (GDPR), and wrongful behaviour of online platforms (Digital Services Act).

16 Rankin, Y. A., Thomas, J. O., & Erete, S. (2021). Black women speak: Examining power, privilege, and identity in CS education. *ACM Transactions on Computing Education (TOCE)*, *21*(4), 1-31. Chakrabarty, D. (2018). 'Provincializing Europe' (*Princeton University Press*, 2007) <https://press.princeton.edu/titles/8507.html> accessed 27 October 2018.

17 Baxi, U. (2012). 'Postcolonial Legality: A Postscript from India' 45 Verfassung und Recht in Übersee / Law and Politics in Africa, Asia and Latin America 178. D'souza, R. (2018). *What's wrong with rights? Social movements, law and liberal imaginations.* Pluto Press.

18 Pahuja, S. (2011). *Decolonising International Law: Development, Economic Growth and the Politics of Uni-*

The chapter introduces a philosophical field of studies which can be referred to under the umbrella terms of epistemic justice and injustice. Work from these fields helps to identify and define injustices in practices of knowledge-making, including law-making, and what strategies to engage in towards prevention and repair. With this we hope to contribute to what we see as very useful bridge-building between the fields of philosophy and law for social justice-oriented scholarship and law-making. The purpose of our chapter is not to offer a solution (i.e. a template for an epistemically-just legislation on AI) but to introduce and promote a possibility for reflection in seeing the law not just as an instrument of regulatory power but as an instrument of epistemic power and in/justice.

Section 2 further introduces the concepts of epistemic in/justice and the case that can be made for engaging with these notions in legal research generally. It relates lawmaking to knowledge-making and demonstrates why understanding this relation in the context of social justice requires explicit attention. Section 3 argues the need for such engagement with the EU's law-making ambitions for the AI Act in particular, among other things by calling attention to the historically problematic precedents of Europe as dominant norm-setter for both law and knowledge practices. Section 4 illustrates what it can mean to approach EU law-making through an epistemic injustice lens by reading parts of the Act's text as recently agreed on by the EU at the time of writing. We end the chapter with a discussion of the value of engaging in this work.

## 2. Law-Making, Knowledge-Making, World-Making

There are many epistemic, or knowledge-related dimensions to lawmaking. We call attention to how lawmakers necessarily choose and work with knowledge(s) about the world and its inhabitants. Law uses such knowledge as an interpretative tool to assume, interpret and inform the making of rules; to decide about (further) types of expertise to accept and refer to; and to determine how people can interact and engage with the law.

Lawmaking is not just a process of rulemaking that requires compliance. It is also the basis upon which our understanding about how the law is used, practiced, interpreted and adjudicated is built. Situating the formation of rules and principles within the political economy in which they are created is critical because this political economy is not neutral, and as Maldonado argues it "presupposes a subject, a space, and a time that determine the way we understand the processes that allow for the emergence, trading, and consumption of legal knowledge".[19]

*versality*. Cambridge University Press.

19 Maldonado, D. B. (2018). The political economy of legal knowledge. In *Constitutionalism in the Americas* (pp. 29-78). Edward Elgar Publishing.

Law creates legal knowledge through defining conceptual categories centring for example crime, trade, property, care, endangerment, rights, and obligations; setting qualifying standards for such terms (knowledge, facts, events and other evidence that gets to 'count as' legally relevant) and defining processes through which this is to be done.[20] As a result, lawmaking has an important function in terms of building knowledge about the societies it regulates, including how lawmaking effects and shapes the worlds it acts in, and what people should expect from this. It is therefore important to recognise the process by which lawmakers codify their understanding of, and ambitions for, a regulated community. Foundational decisions include who counts as part of the community, who has the capacity to participate and be affected by law in this community, and who deserves protection. Think of how the Human Rights regime is primarily organized around the autonomous, individual citizen subject and their claims against nation states. Critical authors point out how, among other features, these are restrictions that de facto embed relations of structural oppression.[21] People are also most systematically affected as communities by local and globalized industry in places where they do not have legal standing against a regime. This requires other legal imaginations around human flourishing and harm, and other allocations of responsibility.[22] We also point to how the typical legal imaginations around antidiscrimination fail to take into account what Kimberlé Crenshaw described as intersectionality, where people's experiences are informed by the interaction of multiple dimensions at the same time, such as race and gender. Her work demonstrated that it is inadequate to deploy either a race or a gender lens without looking at the entanglements that they together produce.[23]

As is sometimes downplayed outside of Critical Legal Studies, lawmaking is an avowedly political practice. Most straightforwardly, laws are made by and through (the enactment of) national and international political institutions such as the UN, the Council of Europe, or the EU. We use the term 'political' in this chapter to emphasise

20 An example: the Dutch Supreme Court recently referred itself to the EU court of Justice, asking whether the Dutch State is allowed to demand that asylum-seekers who claim to be risk of political prosecution prove how their political opinion is fundamental, meaning "particularly important for maintaining his or her identity or conscience," or, at least "deeply rooted in that applicant [so] that, on his or her return to his or her country of origin, he or she could not refrain from manifesting them." The answer was negative. <https://curia.europa.eu/juris/document/document.jsf?text=&docid=277631&pageIndex=0&doclang=NL&mode=req&dir=&occ=first&part=1&cid=483175>.

21 See for example Wright, S. (2003). *International human rights, decolonisation and globalisation: Becoming human*. Routledge. Dembour, M. B. (2006). *Who believes in human rights?: reflections on the European Convention*. Cambridge University Press. Whyte, J. (2019). *The morals of the market: Human rights and the rise of neoliberalism*. Verso Books.

22 Solano, J. L., de Souza, S., Martin, A., & Taylor, L. (2022). Governing data and artificial intelligence for all: models for sustainable and just data governance.

23 Crenshaw, K. W. (2013). Mapping the margins: Intersectionality, identity politics, and violence against women of color. In *The public nature of private violence* (pp. 93-118). Routledge.

how rule-making processes involve negotiation, conflict, compromise and strategic power play as part of human social interaction, especially around questions of authority.[24] Injustices that come with such processes are borne into the legal results. In order to ascertain different forms of rule-making and their potential for oppression, it becomes necessary to tease out analytical standpoints like race, feminism, disability, and colonialism.

We have discussed knowledge as both a source and an outcome of lawmaking, and lawmaking as an inherently political practice. We now call attention to the political nature of *all* kinds of knowledge practices and their outcomes, whether these are understood as social, political, legal, or scientific. The politics of knowledge are borne into laws most obviously whenever law chooses notions and concepts to work with, for instance where it codifies wrongful ideas on race or sex into rules around identification. Knowledge that is already problematic consolidates into enforceable norms this way, and lawmaking arguably performs an additional injustice itself when it does that. The philosophical fields of epistemic justice and injustice help to explain these phenomena and what we mean with 'problematic knowledge.'[25]

Authors from these fields attune to relations of social and informational authority. They analyse how knowledge and knowledge practices can respect or disrespect, single out or make invisible, support or disadvantage persons and communities as epistemic agents: as thinkers, learners and knowers, as social actors and as participants in (or subjects to) decision-making. Central themes in this research include 'authority, credibility, justice, power, dis/trust, and testimony'.[26] These translate into questions around who is afforded authority to speak or testify, who is considered trustworthy and credible, and whose and what kind of expertise is afforded the status of knowledge.

An example of a knowledge-making domain whose 'epistemic trouble' runs deep and regularly makes the news is North-Western[27] medicine. Reports on medicine's failure

24 Bacchi, C. (2012). Why study problematizations? Making politics visible. *Open journal of political science*, *2*(01), 1.

25 With regard to the terms, frequently cited are Fricker's conceptualizations of two particular kinds: testimonial and hermeneutical injustice. These roughly translate to wrongfully denied credibility and the denial of epistemic resources or vocabulary. More descriptions exist such as Epistemic Responsibility in Code, L. (1987). Second persons. *Canadian Journal of Philosophy Supplementary Volume*, *13*, 357-382 more types and ontologies of epistemic injustice have been analyzed and developed such as epistemic oppression in Dotson, K. (2014). Conceptualizing epistemic oppression. *Social Epistemology*, *28*(2), 115-138. Mitova, V. (2020). Explanatory injustice and epistemic agency. *Ethical Theory and Moral Practice*, *23*(5), 707-722.

26 Kidd, I. J., Medina, J., & Pohlhaus, G. (2017). Introduction to the Routledge handbook of epistemic injustice. In *The Routledge handbook of epistemic injustice* (pp. 1-9). Routledge.

27 We use the term somewhat symbolically: as Tuck and Yang argue, the phrase 'North Western' itself can be misleading as it excludes indigenous knowledges in the North-Western hemisphere. Tuck, E. & Yang, W. (2012) 'Decolonization Is Not a Metaphor', Decolonization: Indigeneity, Edu-

to research, recognise and act on disease in, among others, Black, female, indigenous populations, and accounts of harmful experimentation on people who are part of these populations, are expressions of how the field's self-organised and largely self-ruled knowledge-making historically prioritises the white male.[28] This example is illustrative for how different injustices, especially institutional social injustices such as poorer healthcare for marginalized groups, have epistemic dimensions. This chapter is concerned with pointing out such relations. It aligns with those who argue to first consider the standpoints of those who are historically and/or systemically wronged when creating laws.[29] Their experiences need to inform knowledge and practitioners' methods and dispositions so these can become a force for good: to *prevent* 'the making, sharing and perpetuation of oppressive epistemic representations'.[30]

Scheman for example argues for *sustainable* knowledge-making: for norms that 'underwrite practices of inquiry that make it more rather than less likely that others, especially those who are variously marginalized and subordinated, will be able to acquire knowledge in the future', adding that her case can be made as much on epistemic as on social justice grounds.[31] Translated to lawmaking as a knowledge practice and as a practice that enacts knowledge-making rules, sustainable law needs to enable more people to argue their cases, to translate their situations into legal imaginations. To some extent the verbal and open-norm character of law sustains such further development by nature. Think of how judges align with social developments in a regulated society by performing progressive interpretations of legal concepts, such as reinterpreting the term abuse

cation & Society 1, nr. 1 (2012). See also Ndlovu-Gatsheni, S. J. (2018). Introduction: seek ye epistemic freedom first. In *Epistemic freedom in Africa: Deprovincialization and decolonization*. Taylor & Francis.

28 Nordell, J. (2021). The bias that blinds: why some people get dangerously different medical care (The Guardian 21 September 2021 <https://www.theguardian.com/science/2021/sep/21/bias-that-blinds-medical-research-treatment-race-gender-dangerous-disparity>), Carel, H., & Kidd, I. J. (2014). Epistemic injustice in healthcare: a philosophial analysis. *Medicine, Health Care and Philosophy*, *17*, 529-540. Dissecting racial bias in an algorithm used to manage the health of populations <https://www.science.org/doi/10.1126/science.aax2342>. Olson, M. (2017). Females Exposed to Nuclear Radiation Are Far Likelier Than Males to Suffer Harm. <https://www.passblue.com/2017/07/05/females-exposed-to-nuclear-radiation-are-far-likelier-than-males-to-suffer-harm/>).

29 For the case to do this in ethical theory, see Mills, C. W. (2005). "Ideal theory" as ideology. *Hypatia*, *20*(3), 165-183, for an example of how knowledge-making practices need to be in focus to further restorative initiatives around on health disparity see for example Pierce, R. L., Gallifant, J., & Celi, L. A. (2023). Tying Equity To Reimbursements. *Health Affairs Forefront*.

30 As one of us argued elsewhere, a set of 'tandem values' is typically construed in descriptions of (what amounts to) epistemic justice. The first leg (accuracy/due care/competence) focuses on the creation of knowledge on responsible terms and investigative strategies. The second (sincerity, intellectual honesty, trustworthiness) on the sharing of knowledge in responsible terms. De Groot: Care to Explain? A critical epistemic in/justice-based analysis of legal explanation obligations and ideals for 'AI'-infused times (Dissertation Tilburg University, 2023)

31 Scheman, N (2012). Toward a Sustainable Epistemology. *Social Epistemology,* vol 26, issue 3-4

to include non-physical variants.[32] Law can also promote or even force change in expert knowledge practices when their uptake of societal progress lags behind.[33] In the context of the chapter, the quest for sustainability would for example ask of law to promote the collaborative development of computational methods that truly cater to the needs of those whose life it has so far negatively affected – and for law to codify this understanding of AI over understandings that perpetuate false narratives of the technology as an equally potent source for 'good', that is, a neutral technology in itself. The former is what an institute like DAIR seeks to do. It argues that 'research should center the voices and experiences of those most impacted by technology and should be rooted in their communities'.[34] In doing so, it advocates for a more grounded approach that accounts for what the material impacts are of the technology, cautioning for when AI causes more harm than good.

## 3. The EU's Regulation of AI: A Momentum for Epistemic Justice-Oriented Engagement

It is with all this in mind that we are interested to unpack and discuss some legal imaginations around the regulation of AI that are being put forward through the EU AI legislation. We regard the AI Act as a piece of knowledge-making about AI with effects on how AI 'gets to be done' after the Act enters into force. These imaginations pertain to ideas about what AI can do, how it works, and what the reasons and rationales for the EU's need to regulate it are.[35] Our argument is that the ways in which knowledge about AI, and about AI as part of knowledge-making practices, are framed in the Act, co-determine conceptual, technical as well as infrastructural aspects of AI.[36] As we will discuss,

32 Hattenstone, S. (2020). 'Lady Hale: "My Desert Island Judgments? Number One Would Probably Be the Prorogation Case"' (The Guardian, 11 January 2020, sec. Law. <https://www. theguardian. com/law/2020/jan/11/lady-hale-desert-island-judgments-prorogation-casesimon- hattenstone>.)

33 For example, legal force was needed to force the uptake in practice of informed consent developments, and we now see various laws pushing for explainable automation in one way or another

34 'About- DAIR Institute' <https://www.dair-institute.org/about/> accessed 21 December 2023.

35 See generally for discussions on imaginations around AI Now Institute, 'A New AI Lexicon: Responses and Challenges to the Critical AI Discourse' (A New AI Lexicon, 19 January 2021) <https://medium. com/a-new-ai-lexicon/a-new-ai-lexicon-responses-and-challenges-to-the-critical-ai-discourse-f2275989fa62> accessed 20 March 2022. Cave, S. Dihal, K. & Dillon, S. (2020). 'Introduction: Imagining AI' in Stephen Cave, Kanta Dihal and Sarah Dillon (eds), AI Narratives: A History of Imaginative Thinking about Intelligent Machines (Oxford University Press 2020) <https://doi.org/10.1093/oso/9780198846666.003.0001> accessed 23 April 2023.

36 Causevic, A. & Sengupta, A. (2020). 'Whose Knowledge Is Online? Practices of Epistemic Justice for a Digital New Deal – A Digital New Deal' <https://projects.itforchange.net/digital-new-deal/2020/10/30/whose-knowledge-is-online-practices-of-epistemic-justice-for-a-digital-new-deal/> accessed 23 April 2023. Critical Data Scholars also need to be careful of their own extractive

this already starts with the definition of AI and with 'AI systems' as the object of regulation. In other words, (the) law's knowledge-making about AI becomes part of what AI is for those on whose lives the law exerts an influence.

Much legal scholarship focuses on the 'how' of automated decision support systems (ADS). The increasing processing complexity of ADS is seen to challenge legal paradigms such as accountability, justification, and reason-giving, all of which rely on decisional insightfulness. Scholarly problematisations abound with regard to AI's lack of understandability and concerning the adequacy and appropriateness of legal responses in light of their analyses.[37] We will not summarise this research here, but want to point out that where some research is more concerned with 'saving' regulation as we know it,[38] others point out how ADS challenges reveal the moral and ethical inadequacy of traditional legal frameworks in ways mentioned in the previous section (such as the individualist frame). This inadequacy is also identified as a flaw with the majority of the law-tech field, and as was cited, the field can be seen to respond by accepting its need for reflection and change.[39]

This ties in with the increasing recognition that so many voices are still excluded from regulatory institutions as well as the computational industry. For a perverse expression of this phenomenon, we can refer to a recent call of several EU countries on the EU high representative Josep Borrell in the context of the EU's role in ending global inequality. They asked him to 'coordinate a diplomatic campaign to woo countries that have felt ignored', citing the need to do a better job of competing with China and Russia in what Borrell has described as a 'battle of narratives'.[40]

There are many powers pulling at law to move it in a certain direction and this includes the AI industry. In the runup to the finalisation of the AI Act, they too are involved in battles over how AI is defined and what makes its applications safe or risky.[41] The out-

actions, Lehuedé, S. (2024). The double helix of data extraction: Radicalising reflexivity in critical data studies. *Technology and Regulation*, 84-92.

37 See e.g. Maranke Wieringa on accountability challenges: What to account for when accounting for algorithms: a systematic literature review on algorithmic accountability (FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency January 2020 <https://doi.org/10.1145/3351095.3372833>), Kaminski, M. (2019). The Right to Explanation, Explained (Berkeley Technology Law Journal, Vol. 34, No. 1, 2019), and on consumer protection law Sax, M., Helberger, N., & Bol, N. (2018). Health as a means towards profitable ends: mHealth apps, user autonomy, and unfair commercial practices. *Journal of consumer policy*, *41*, 103-134).

38 Brownsword, R. (2019). Law disrupted, law re-imagined, law re-invented. *Technology and Regulation*, *2019*, 10-30.

39 See note 12.

40 Context: Guardian Article 18 Sept 2023 "Western leaders defend slow progress to end global inequality as UN summit starts" <https://www.theguardian.com/world/2023/sep/18/un-summit-sees-western-leaders-defend-slow-progress-to-end-global-inequality>.

41 See for example: Connor Axiotes: Lobbying for Loopholes: The Battle Over Foundation Models in the EU AI Act (<https://www.euractiv.com/section/digital/opinion/lobbying-for-loopholes-the-battle-over-foundation-models-in-the-eu-ai-act/>), Amnesty International: EU: France, Germany

comes will matter for the protection against AI's harms. This is all the more reason to make use of the momentum that exists for regulation and governance to ensure that excluded voices are part of the processes that influence regulation. Bringing in different perspectives will help with understanding the impacts of AI and law, and also bring a repository of knowledge that helps to understand and build more effective solutions.[42] For instance, thinking with a decolonial lens about AI helps to ascertain how it continues to further colonialism.[43] This can be seen in relation to the use of land and resources for creating data centres, in terms of how power over development remains in the hands of a limited number of big corporations, and in their influence on how AI is regulated.[44]

In this battle over the AI narrative that the EU also claims it has a globally beneficial role to play in, the EU is the self-acclaimed global centre of value-driven, fundamental rights-protecting technology regulation.[45] The objective of the AI Act is to ensure that the 'Union is global leader in the development of secure, trustworthy and ethical artificial intelligence' and to 'ensure the protection of ethical principles'.[46] The abundant mentions of values and ethics create an expectation for the 'imagining' of AI that departs from the thus far adopted narrative in EU law of technologies as value-neutral rather than inherently value-laden,[47] and of AI as a scientific paradigm that can safely be asked to self-regulate its beneficial way forward. The centring of the EU as global value leader, on the other hand, raises concerns about the Act's restricted assignment of institutional capacity to particular geographical or political economies based on their capabilities, independence, transparency, accountability, and capacity to monitor AI. Put differently, whose knowledge and expertise about AI is prioritised in the Act? Relatedly, are Euro-

and Italy risk unravelling landmark AI Act negotiations (<https://www.amnesty.org/en/latest/news/2023/11/eu-france-germany-and-italy-risk-unravelling-landmark-ai-act-negotiations/>).

42 Nyabola, N. (2024). Ngugi and Mazrui in Digitalization Policy: Practitioner Insights into The Role of Language in Decolonising Digitalisation Policy. *Technology and Regulation*, *2024*, 63-72.

43 Mittal, A. (2024). Constitutionalism as a Way to Decolonize Global Data Law Development. *Technology and Regulation*, *2024*, 19-27. Mery, V. (2024). The Chilean constitutional-making process: a case study in decolonising and reframing digital governance. *Technology and Regulation*, *2024*, 103-114.

44 See for example Katz, Y. (2020). *Artificial whiteness: Politics and ideology in artificial intelligence*. Columbia University Press. de Souza, S. P., Smith, H. M., & Taylor, L. (2024). Decolonial Data Law and Governance. *Technology and Regulation*, *2024*, 1-11. Mesquita, H., Garrote, M. G., & Zanatta, R. A. (2024). Regulating Artificial Intelligence in Brazil: the contributions of critical social theory to rethink principles. *Technology and Regulation*, *2024*, 73-83. Lulz, S. D. (2024). Tactics of Earthy Data: Decolonising for the Anthropocene. *Technology and Regulation*. Mishra, K. (2024). Data as a national asset: What does seeing data in terms of an asset reveal about postcolonial state in India?

45 Boshe, P., & Caride, C. G. (2024). Is the Brussels Effect Creating a New Legal Order in Africa and Latin America and the Caribbean?. *Technology and Regulation*, *2024*, 12-18.

46 Recital 5.

47 Campolo, A., & Crawford, K. (2020). Enchanted determinism: Power without responsibility in artificial intelligence. *Engaging Science, Technology, and Society*.

pean values posited as uncontested, or does the Act show reflection on the critique that these values and the technology law field have received?[48]

The EU plays an important and arguably outsized role as the de facto regulator of the world that is not restricted to technology-oriented law.[49] Its legislative instruments determine matters from personal data protection and environmental protection to trade and business. Its effects, as Bradford for example argued, assume global importance on account of the strength of its market, its regulatory capacity, stringent standards, institutional capacity, and non-divisibility.[50]

As a result, whether by intentional design through the extra-territorial application of its law, or by consequence of its regulatory power, countries and economies across the world are required to comply with its rules and regulations to be able to trade and do business with the Union. Underlying its regulatory influence and power, which is an argument that is not new, there is an epistemic power that the EU exerts over the world by virtue of the ways in which it describes, defines, and categorises different phenomena, and the rules by which these must be understood.[51] This is the knowledge-making power that highlights that certain practices, beliefs and customs have greater prominence and visibility over others.[52] There is a divide that promotes certain types of knowledge to the detriment of others.[53] It creates distinctions where other knowledge is deemed to have less value, less clarity and capability to be able to provide comparable alternatives.[54]

We want to highlight that the creation of regulation for and by the EU must also be analysed in terms of the consequences it will have for the legal imaginations in terms of institutions, procedures, and substance. Doing so recognises the power imbalances that shape how regulation is made, where often, conceptual categories in the Global North trump those of the Global South. This results in a transplantation of ideas and creates a hierarchy of norms and values, without a robust analysis of their attendant consequences. These critiques are urgently relevant in light of law's failure to protect people from technology-fuelled harms that manifest most prominently on the institutional and systemic levels that are less adequately addressed by law.

48 Solano, J. L., de Souza, S., Martin, A., & Taylor, L. (2022). Governing data and artificial intelligence for all: models for sustainable and just data governance.

49 Outsized in relation to territory and geographical scope.

50 Bradford, A. (2020). *The Brussels effect: How the European Union rules the world*. Oxford University Press, USA.

51 Dotson, K. (2014). Conceptualizing epistemic oppression. *Social Epistemology*, *28*(2), 115-138.

52 Eslava, L., & Pahuja, S. (2012). Beyond the (post) colonial: TWAIL and the everyday life of international law. *Verfassung und Recht in Übersee/Law and Politics in Africa, Asia and Latin America*, 195-221. Raghunath, P. (2024). Critical data governance: A southern standpoint to the study and practice of data. *Technology and Regulation*, *2024*, 37-46.

53 Anzaldúa, G. (2007). *Borderlands: The new mestiza*. Aunt Lute Books.

54 Tuhiwai Smith, L. (2012). *Decolonizing methodologies: Research and indigenous peoples*. Zed books.

## 4. Reading (Parts of) the EU AI Act: A Reflective Exercise

### 4.1 Introduction: The Act's Problematisation of AI

The previous sections articulated several directions for a sustainable[55] approach to lawmaking-as-knowledge-making about AI. To reiterate, we understand AI as a socio-technical practice that has problematic ideological roots. AI is used for creating actionable outputs, presented as 'knowledge' or 'intelligence'. Lawmaking about AI ideally at the very least enables affected persons to recognise their situations in the legal terms so that they can argue their cases and redress any grievances. An objective of law should be to force necessary change in the typically self-regulating technological practices towards justice-oriented reform.[56] Thus law around AI should support a grounded, critical understanding of AI over hyped and obfuscating claims, and sustain an inclusive and critical further development of the regulated paradigm. In the case of the EU as lawmaker specifically, the territory carries a weighted responsibility to get this right. It must not only regulate a large and important market, but needs to ensure that its regulation does not perpetuate an exclusionary and oppressive agenda in the effort, mindful of the region's historical record both with regard to lawmaking, knowledge-making, world-making before and after colonial times, and of its extraterritorial effects today.

With these cues in mind, this section offers a set of reflections on the legal imaginations about AI and its regulation that the EU lawmaker agreed on. The reflections offer illustrative points for those working on or with the regulated AI landscape in any capacity in the wake of the Act's entry into force.

The reflections are based on the latest available texts of the EU's 'Proposal for a Regulation laying down harmonised rules on artificial intelligence'.[57] Although several provisions in the Act have been studied, a large part of the illustrations are sourced from the Act's recitals. Underlying this choice is how the recitals, although not legally binding of themselves, elaborate authoritatively on the why's and how's of the provisions laid down in this 'landmark law'. Recitals state the reasons for enacting rules in a particular way, and the objectives they pursue. They translate the EU's values and principles into the law-specific domain; they are used by the Court of Justice of the EU (CJEU) in their

55 Building on Scheman, section 2.

56 For example, require truly inclusive research & development processes, informed by the experience of those who've borne the brunt of AI mishaps

57 We base ourselves on two documents: the 'compromise text', i.e. the text that was voted on when the AI Act was adopted in March 2024, published by the Council of the EU in the document dated 26 January 2024, <https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf> and the document that also includes previous versions of the Act's proposal in a four-column overview, Posted by journalist Luca Bertuzzi on LinkedIn, <https://www.linkedin.com/posts/luca-bertuzzi-186729130_aiactfinalfour-column21012024pdf-activity-7155091883872964608-L4Dn/>.

putative interpretation of EU law, as well as by those working with EU law in other capacities.[58] Their non-literal binding status allows recitals to be drafted in a more narrative way, more easily revealing of the law's politics than can be gleaned from provisions alone.

Finally, a word on the term 'problematisation'. We use it here in the way that Bacchi for example uses it in her method for studying policy documents: to find out how the policy subjects and objectives are performatively[59] defined, to advance how both are understood and pursued and to encourage aspirations that are tied to these understandings.

### 4.2 Framing AI

Recital 3 of the Act describes AI as 'a fast evolving family of technologies that contributes to a wide array of economic, environmental and societal benefits across the entire spectrum of industries and social activities'. The word 'can' preceded 'contribute' in earlier drafts; its absence now supports a decidedly techno-optimist framing. That tone is softened further on where the recital states how AI *can* support 'socially and environmentally beneficial outcomes'. The remark is still followed by a rather ambitious list of example domains, seemingly illustrating the 'entire spectrum of industries and social activities'– including protections from human-induced climate disaster.[60]

This codified optimism has the potential to create lasting effects. Most obviously the Act will be a factor in what the AI industry will be allowed to try and accomplish. In a positive reading, this is the EU lawmaker promoting the kinds of AI practices that the world does not yet see enough of. But in light of what we know about the harms that result from how AI is done, idealist expectations that do not start from this 'trouble' arguably codify their continuation.[61]

The AI-induced 'dangers' that the world *has* seen enough of are mentioned in the shorter recital 4. These are referred to in terms of potential rather than fact. Their mate-

58 Recital 71 of the EU's General Data Protection Regulation is a case in point, explaining the purpose of granting a right to explanation which is not explicitly stated in provisions. The recital is used by the EDPB in their authoritative interpretation (Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 (wp251rev.01)

59 On this point specifically see also the application of Ulnicane, I., Knight, W., Leach, T., Stahl, B. C., & Wanjiku, W. G. (2021). Framing governance for a contested emerging technology: insights from AI policy. *Policy and Society*, *40*(2), 158-177. Bacchi, C. (2012). Why study problematizations? Making politics visible. *Open journal of political science*, *2*(01), 1.

60 "Healthcare, farming, food safety, education and training, media, sports, culture, infrastructure management, energy, transport and logistics, public services, security, justice, resource and energy efficiency, environmental monitoring, the conservation and restoration of biodiversity and ecosystems and climate change mitigation and adaptation." Recital 3, Compromise text 26 Jan 2024

61 Mills, C. W. (2005). "Ideal theory" as ideology. *Hypatia*, *20*(3), 165-183.

rialisation is made out to be dependent on use rather than design and continuous development,[62] and, crucially, on technological advancement.[63] Put differently, AI's *envisioned* beneficial potential is prioritized over AI's *experienced* harms. This is an epistemic injustice in itself. The fact that no applicable legal and ethical frameworks to date have meaningfully stood in the way of the manifestation of 'data-driven' harms deserves an account that speaks to the knowledge and experiential positions of those affected: an account that engages with how this is so, and how the Act stands to make a difference.

As mentioned in previous sections, one reason that legal protections have failed is their organisation around individual harms and remedies, a disadvantage that is exacerbated in light of AI's systemic influence. Yet the Act makes a point of describing harms as individual, at most admitting these harms may affect large groups of similar persons.[64] The Act especially focuses on a particular kind of individual harm labelled *mis*use, with a warning that AI can 'provide novel and powerful tools for manipulative, exploitative and social control practices'.[65] Recital 16 zooms in on subliminal techniques (described as unperceived, unable to resist) that thwart behaviour and nudges persons in a way that 'subverts and impairs their autonomy, decision-making and free choices'. But in light of AI's extant harms, the rather intended 'nudging' of decision makers who (are made to) work with computational systems that automate marginalisation and adversarial treatment of groups of people is arguably a much more pressing problem.[66]

Related to the previous point is the Act's decision to exclude from the framing of AI and why it needs regulation, those technologies that more straightforwardly do what they are instructed to do by humans. The Act prioritises AI's perceived capacity to 'infer' for example 'predictions, content, recommendations, or decisions'. It defines such activity as more complex and therefore 'riskier'. It chooses to ignore harms that follow from less technologically complex applications, which ties in with dominant AI industry framings that solely see dangers in what they define as 'cutting edge'. This is a danger-

62 As studies have demonstrated again, AI builds on human labour which is underpaid, exploitative, and even violent to e.g. those who have to exercise supposedly mundane tasks of coding, correcting, and cleaning data. Just two examples from different continents: on data labelling in Venezuela, <https://www.technologyreview.com/2022/04/20/1050392/ai-industry-appen-scale-data-labels/>, on GPT cleaning in Kenya <https://www.wsj.com/articles/chatgpt-openai-content-abusive-sexually-explicit-harassment-kenya-workers-on-human-workers-cf191483>.

63 The emphasis on 'advanced' can be read as an acceptance of big tech's claims with regard to what is and isn't dangerous and who to turn to for solutions.

64 For example, recital 18 accedes that remote biometric identification "may affect the private life of a large part of the population."

65 Recital 15

66 Which of course certainly also leads to autonomy harms for those who are affected by e.g., automated policing, hiring practices, visa applications, border control to name just a few examples. Hannah Arendt's work imposes itself when making this point: Arendt, H. (1994). Some questions of moral philosophy. *Social research*, 739-764.

ous distraction[67] that downplays the convoluted and influential ways that 'simple' algorithms and AI have become entangled in the establishment of harm in human decision-making practices.[68]

The narrative is reinforced at various points where specific AI activity in the context of a high-risk category is exempted from the regulation. Examples are 'preparatory' work and work done to 'improve the result of a previously completed human activity'.[69] The Act offers a surprisingly specific example of the AI-driven identification of a teacher who digresses from their own previously established grading pattern 'so as to flag potential inconsistencies or anomalies'. Different arguments can be made just as well. In the context of education, where biased assessment and assessment tools as well as disparate quality of pre/primary education are established problems, deviation from grading patterns may be precisely what is necessary to address systemic inequalities.[70]

Another instance of the Act's framing of envisioned beneficence and extant harm, is the way in which AI-driven 'climate change mitigation and adaptation' are heralded. This does not account for the fact that data centres around the world are being criticized for being among the biggest guzzlers of water and electricity, rendering great difficulty for people to access such resources.[71] Such siloed analysis of the effects of technology offers the possibility for cementing the techno-optimist outlook presented earlier.

Before moving on to several reflections on risk, AI literacy, and the recitals' address of what it terms 'pre market' AI activity, such as research, development, and testing, we add a few words on 'systems' in the Act's performative framing of AI. The Act only applies to AI systems, described as software products that run on machines, either stand-alone or as components of other products.[72] The context of a system is its 'environment', that

67 See footnote 8 and <https://managing-ai-risks.com/>. See also Crawford and Joler who explore the implications of the Amazon Echo system in the realms of human labour, data, and planetary resources: "The full stack [required to interact with Echo] reaches much further into capital, labour and nature, and demands an enormous amount of each." <https://anatomyof.ai/>.

68 See, e.g., Saxena, D., Repaci, C., Sage, M. D., & Guha, S. (2022, April). How to train a (bad) algorithmic caseworker: A quantitative deconstruction of risk assessments in child welfare. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts* (pp. 1-7). Binns, R., & Veale, M. (2021). Is that your final decision? Multi-stage profiling, selective effects, and Article 22 of the GDPR. *International Data Privacy Law*, *11*(4), 319-332.

69 Recital 32a.

70 Kotzee, B. (2017). Education and epistemic injustice. In *The Routledge handbook of epistemic injustice* (pp. 324-335). Routledge. Later on, the Act acknowledges how AI can be a force for worse in education: recital 35 states that (educational) systems can be "improperly designed and used" to e.g. perpetuate discriminatory patterns.

71 Another instance of the Act's framing of envisioned beneficence and extant harm, is the way in which AI-driven "climate change mitigation and adaptation" are heralded. This does not account for the fact that Data Centers around the world are being criticized for being among the biggest guzzlers of water, and electricity, rendering great difficulty for people to access such resources

72 Recital 6, Article 3

which a system influences. The terminology expresses a choice against understandings put forward in studies across historical, technological, and other disciplines that describe 'AI systems' as dynamic amalgams of technologies, people, human and digital networks, power relations and practices, and instead defines AI as a term that at any time carries the meaning that key players want it to have.[73] The EU is one of these players, as the European Commission is tasked with authority over the explanation and application of the Act's definition of 'systems'. A special AI Board and AI Office are also foreseen with interpretation and application tasks. Envisioned engagement with 'the' scientific community is named at various moments, unqualified with regard to discipline.[74] In light of how the Act deprioritizes social-critical understandings of AI and that of critical technology communities, we should wonder who will be part of that scene. In any case, all these institutions will become heavily-lobbied sites.

### 4.3 Risks Not Rights

The AI Act's approach to developing binding rules draws from a risk-based approach. Recital 14 states that:

[t]hat approach should tailor the type and content of such rules to the intensity and scope of the risks that AI systems can generate. It is therefore necessary to prohibit certain unacceptable artificial intelligence practices, to lay down requirements for high-risk AI systems and obligations for the relevant operators, and to lay down transparency obligations for certain AI systems. [75]

In creating a hierarchy of risks, the Act undertakes to establish red lines for conduct and confers the power to determine what is acceptable or not as well as who has the knowledge to do so to the law and law maker. The Act therewith claims an expert role for law and its imagined institutions with regard to framing an important (technical, social, political) kind of knowledge. The Act's 'hierarchies of risk' prescribe knowledge about the effects of AI and suggestions on how it should be governed.

Article 6(3) of the draft agreement provides for exceptions to high-risk classifications where it is said not to pose a risk to a person's health, well-being, or fundamental rights.[76]

73 See for example the earlier cites authors (Katz, Benjamin, Lepore, Broussard), and critical reports such as Balayn and Gürses: Balayn, A., & Gürses, S. (2021). Beyond Debiasing: Regulating AI and its inequalities. *EDRi report*.

74 Recital 60n, 75; *Main elements of the compromise*," point 9

75 Recital 14.

76 Article 6(3) "an AI system shall not be considered to be high-risk if it does 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. This shall be the case where one or more of the following conditions are fulfilled: (a) the AI system is intended to perform a narrow procedural task; (b) the AI system is intended to improve the result of a previously completed human

The Act provides procedures by which such an application for assessment can be made.[77] It states that given the complexity of the nature of the technology, for high-risk products, in the early phase of the regulation, such risk assessment should be carried out by the providers themselves.[78] This framing does not just invest providers with responsibility but also with a high degree of trust in the outcomes of their assessments. In this way, the law identifies who has the capacity to give testimony on behalf of society on the effects that may emerge on account of the technology.

When providers apply for an exception under Article 6(3), the Act provides methods to have the veracity of such an application evaluated by a market surveillance authority that is also a national authority. This authority is deemed to have the competency as well as the necessary powers to be able to obtain data to make their assessment.[79] It is worth noting that such a provision presumes that such bodies will have a sufficient number of staff and the necessary personnel, as well as scientific qualifications to undertake such work across national jurisdictions. In addition, given the distributed nature of such bodies across countries, it is also necessary to evaluate whether such rigour can be maintained across the Union, with different resources and capacities available to union members.

Finally, Article 85 provides any natural or legal person, without prejudice, can submit a 'reasoned complaint' to the market surveillance authority.[80] In articulating this process, the act demonstrates that the assessment of what is high risk, is as much a matter of scientific assessment as it is an interpretative exercise, where the stakeholders involved in such a process are those with the power to determine the nature of the risk. Whereas providers and authorities are assumed to have the necessary skill set to engage in assessments of risks, complaints made by natural or legal persons, are prefixed with reasonability, attributing a qualifier to the nature with which they may express a grievance. In doing so, the act qualifies the nature of testimonies that can be made by different stakeholders, an epistemic injustice by design.

At the heart of a risk-based approach is also the deprioritisation of rights, and importantly, therewith of obligations that could usefully cater to such rights.[81] The choice that the act makes is in regard to a variety of obligations that accrue to developers. In some instances, these are transparency obligations, and in others, they may also include

activity; (c) the AI system is intended to detect decision-making patterns or deviations from prior decision-making patterns and is not meant to replace or influence the previously completed human assessment, without proper human review; or (d) the AI system is intended to perform a preparatory task to an assessment relevant for the purposes of the use cases listed in Annex III.

77 Article 49 provides for registration.

78 Recital 64.

79 Article 70.

80 Article 85.

81 Moyn, S. (2018). *Not enough: Human rights in an unequal world*. Harvard University Press.

human rights risk assessments. But as civil society organisations have long argued, rights are 'not negotiable' and need to be respected regardless of a company's obligations, or the potential risk exposure.[82] By not taking a rights-based approach, the act therefore ends up having to determine and regulate spaces of exceptions/exemptions. For instance, in the case of biometric ID, exceptions exist for law enforcement in terms of emotion recognition as well as remote use of biometric ID. This is despite years of research from organisations across Europe showing how these technologies are error prone and can cause enormous harm.[83] Article 49, which talks of registration in relation to high-risk systems also puts in place exceptions for law enforcement, migration, asylum, and border control from making their registration in the EU Database of high-risk systems public. In doing so, they take away the possibility of wider publics engaging with the work of these entities, and interrogating their uses.

In terms of the fundamental rights impact assessment mechanism provided in Article 27, the Act states that it must be carried out by deployers of high-risk systems. These deployers could be entities which are governed by public law or private entities. Crucially, the conditions of such assessment include aspects related to the processes in place to oversee the intended purposes, the time period of use, the natural persons who would be impacted, the mechanisms of human oversight available, as well as mechanisms for complaints and internal governance.[84]

At the face of it, however, the impact assessments are conducted by the developer itself, allowing for the framing of risk as something that must be understood in line with the intentions of the technology, rather than emerging from the concerns of affected populations. This top-down approach removes the capacity of affected populations of being able to articulate their concerns. Whereas the framing of the assessment names the addressing of individuals and groups, these too are identified by the developer itself. This might also be in contradistinction with the act, which recognises that the nature

82 Hidvegi, F., Leufer, D., & Masse, E. (2021). The EU should regulate AI on the basis of rights, not risks. *Access Now*, *17*. <https://www.accessnow.org/eu-regulation-ai-risk-based-approach/> accessed 22 March 2022

83 Shubham, 'EU AI Act Will Fail Commitment to Ban Biometric Mass Surveillance (Deutsche Version Unten)' (*Reclaim Your Face*, 18 January 2024) <https://reclaimyourface.eu/eu-ai-act-will-fail-commitment-to-ban-biometric-mass-surveillance/> accessed 21 April 2024.

84 Article 27, 'a description of the deployer's processes in which the high-risk AI system will be used in line with its intended purpose', a description of the period of time within which, and the frequency with which, each high-risk AI system is intended to be used; the categories of natural persons and groups likely to be affected by its use in the specific context the specific risks of harm likely to have an impact on the categories of persons or groups of persons identified pursuant point (c) of this paragraph, taking into account the information given by the provider pursuant to Article 13; description of the implementation of human oversight measures, according to the instructions for use, the measures to be taken where those risks materialise, including the arrangements for internal governance and complaint mechanisms'

of discrimination is likely to be intersectional in nature, and that the identities of groups themselves are likely to be fluid and dynamic.[85]

The addition of a fundamental rights impact assessment framework must also be understood in the wider context of the legislation itself, where rights are introduced in a manner where they are balanced with the need to promote innovation, strengthen the internal market, and promote the free movement of AI goods and services.[86] In undertaking a balancing between competing purposes, the Act also creates the possibility that such an impact assessment may become a technical or standard setting exercise, rather than an engagement with considerations of exclusion or violence that emerges for a variety of reasons, including discrimination mentioned in the Act, such as on the basis of gender, race, etc.

The problem of balancing rights, rather than viewing them as intrinsic protections, is witnessed in the manner in which the EU considers the protections of the rights of people who are outside of the EU from technologies created within the EU. The Act does not provide guidance on the export of harmful technologies developed in the EU. As a network of NGOs have argued, it will still be possible to export a system banned in Europe 'despite existing evidence of human rights violations facilitated by surveillance technologies developed in the EU in third countries (e.g. China, Occupied Palestinian Territories).'[87]

This racialised nature of rights protection is also evidenced by concerns raised by civil society where technologies such as risk assessments or predictive technologies are not banned in migration contexts. For instance, the Act explicitly deprioritises the safety of people fleeing areas that are rendered unliveable by the climate change produced by earlier 'competitive advantages' of the Global North's industry. The Act provides little concern for their safety vis-à-vis that of EU citizens in how it mostly excludes the immigration and asylum domains from the Act's application. Evidently, the EU borders will remain a testing ground for harmful AI.[88] This way the Act reserves a dubious role for AI in problematising the social reality of the climate emergency. In fact, NGOs have also argued how EU databases concerned with migration, such as the Eurodac, the Schengen

85 For instance, recital states 18 that the "Technical inaccuracies of AI systems intended for the remote biometric identification of natural persons can lead to biased results and entail discriminatory effects. This is particularly relevant when it comes to age, ethnicity, race, sex or disabilities."

86 Recital 1.

87 <https://www.accessnow.org/press-release/joint-statement-ai-act-fails-migrants-and-people-on-the-move/>.

88 Borak, M. (2024). 'Is the EU AI Act Leaving a Backdoor for Emotion Recognition? | Biometric Update' (16 February 2024) <https://www.biometricupdate.com/202402/is-the-eu-ai-act-leaving-a-backdoor-for-emotion-recognition> accessed 21 April 2024.

Information System, and the ETIAS, will have not be required to comply with the AI Act until 2030. [89]

### 4.4 Act-ing in Support of Fabrication: The Case of Emotion Recognition

The Act includes an inconsistent problematisation of 'emotion recognition', part of a controversial practice of technological developments with a notable lack of scientific grounding.[90] The applications carry a high risk of racist and otherwise discriminatory output[91] and know especially dubious test cases and use outside EU borders.[92] This shorter section connects the previous discussion of the risk-approach to the section after this one, which will engage with the Act's framing of science, research, and development.

Rather than red lining emotion recognition as a misleading practice, the Act codifies the technology by acknowledging its potential high-risk impact, creating specific rules for it, and downplaying the lack of scientific grounding for it. This effectively grants the technology industry permission to simply continue as they are, by only making their practices 'safe,' insofar that the applications can avoid selective prohibitions and safety demands. Recitals 55 and 8 state that biometric data can allow for the recognition, identification or inference of emotions of natural persons. The equally problematic 'reading' of sexual orientation, personality traits and personal attachments to religion and philosophical beliefs, for example, is to some extent validated in recitals 7b and 16a, which add several categories to which persons can be assigned based on their biometrics.[93] Recital 8 again attaches the risk of harm of the applications to their technological advancement, and creates a scientifically untenable difference between detection and inference. 'Mere' detection of 'readily apparent expressions, gestures or movements' like frowns and smiles are excluded from regulation, unless this is attempted with the aim 'to infer emotions'.

The controversial scientific status of the governed systems are referred to as 'concerns' in recital 26b. The recital particularises concerns to the reason that emotions 'vary considerably across cultures and situations, and even within a single individual'. This

89 Joint Statement – A Dangerous Precedent: How the EU AI Act Fails Migrants and People on the Move' (*Access Now*) <https://www.accessnow.org/press-release/joint-statement-ai-act-fails-migrants-and-people-on-the-move/> accessed 21 April 2024. <https://www.accessnow.org/press-release/joint-statement-ai-act-fails-migrants-and-people-on-the-move/>.

90 Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological science in the public interest*, *20*(1), 1-68.

91 Rhue, L. (2018). Racial influence on automated perceptions of emotions. *Available at SSRN 3281765*.

92 Amnesty International, Primer: Defending the rights of refugees and migrants in the digital age.

93 Also named are more usual labels such as age, sex, and race, whereby both 'sex' and 'race' are named without reference to their controversial statuses.

does not do justice to the much more fundamental problems with biometric emotion categorisation and recognition science, as the problem is basically presented as one of insufficient data here.

### 4.5 Batteries Not Included

Principally excluded from direct regulation are key features of what makes AI work, referred to by the Act as pre-market: design, research, development, testing (hereafter: R&D). The exceptional position of these features is expressed at several points in the Act, like the statement in recital 12c that the Act 'should support innovation, respect freedom of science, and should not undermine research and development activity', as well as more literally in the exclusion of models and systems 'specifically developed and put into service for the sole purpose of scientific research and development'[94] and the exclusion of 'product-oriented research, testing and development activity' from the Act's application.

This framing of R&D paints a 'before and after', turnkey picture of how AI is established as a world-affecting force. It also divides the world in producers and consumers. Both decisions obfuscate how affected 'consumers' are an active part of AI's establishment. Their frequently exploited contributions come in the form of moderation, extracted or ignored bodies of knowledge and labour, of captured traces of in/activities, in the form of interactions with digital interfaces across contexts and jurisdictions, and so on. Consider also how the prototyping of 'general-purpose AI models with systemic risk' is added to activities excluded from the Act; this is a choice that expresses blatant alliances with 'big tech's' interests in face of ample evidence to act contrarily.[95]

At several points later on, the recitals backtrack on this framing. The need for some modality of R&D regulation is acknowledged, but the terms remain unclear. Recital 12c states how R&D activities need to comply with ethical and professional standards of scientific research and applicable EU Law 'under all circumstances'. Yet what counts as 'scientific research' in AI development is not self-evident. The label is interpreted creatively to capitalise on industry opportunities and avoid scientific domain regulations in place, including research ethics.[96]

94 Recital 12c.

95 Despite the mess that was triggered by the blunt release of several GPAI's such as Chat GPT, the additional statement that compliance with an array of governance requirements "can be achieved through codes of practice, which will be developed by the industry" can also be seen in this light. *Main elements of the compromise*," point 8 on p. 4/5.

96 See e.g., Sharon, T. (2016). The Googlization of health research: from disruptive innovation to disruptive ethics. *Personalized medicine*, *13*(6), 563-574.

The reference to scientific standards is arguably also inadequate that no mention is made of how these frameworks have not prevented salient harms to people before or after AI became a factor. This is especially weak vis-à-vis the R&D governance needs in high stakes environments like medicine/healthcare: the knowledge-making domain that produced the most influential research standards, whose predicted profits from and expectations for AI are sky-high, and who are still struggling to become more just, equitable and fair as was discussed before. Baumgartner et al. describe in detail which governance needs for healthcare R&D are crucial to ensure so that AI does not continue to exacerbate or simply maintain the existing issues.[97] After all, it is evident that the lack of robust standard setting is all the more harmful in light of how the Act's benchmarks and the technologies developed under it stand to be transferred to regions with less regulatory capacity.[98]

Recital 72c seems to come closest to acknowledging some of the cited needs, but with a twist: to ensure beneficial outcomes for AI, Member States (MS) are 'encouraged to support and promote' AI for social and environmental *problem solving*. The Act expects these projects to be grounded on 'the principle of interdisciplinary cooperation between AI developers, experts on inequality and non-discrimination, accessibility, consumer, environmental, and digital rights, as well as academics'. Similar engagement with the excluded R&D spheres can be found recital 60's statement that 'representatives of groups of persons likely to be affected by the AI system, independent experts, and civil society organisations' *could* be involved in fundamental rights impact assessments and in 'designing measures to be taken in the case of materialization of the risks'. Although the inclusion of others deserves endorsement, this also responsibilises 'consumers' whereas their work would be more worthwhile if they were invited and awarded as 'producers' of technologies with active capacities to influence its design.

We also consider the engagement with the 2019 Ethics Guidelines for Trustworthy AI in recital 14a.[99] Although the referral is long and elaborate, the recital emphasizes the 'optional' status of the guidelines: 'seven non-binding ethical principles for AI which should help ensure that AI is trustworthy and ethically sound'. Recalling the fierce critique on salient weak spots of the principles in places where industry pressure leaned in,[100] the principles could have at least used a firmer endorsement. Instead, after running

97 E.g., a plurality of disciplines and actors, properly supported participatory research activities with historically marginalized methods, reflection, awareness raising and engagement with AI developers norms, assumptions and beliefs "regarding diversity, intersectionality, and justice to understand how these may shape the design and implementation of new technologies" Fair and equitable AI in biomedical research and healthcare: Social science perspectives, Renate Baumgartner et al, Intelligence In Medicine 144 (2023) 102658

98 WHO guidance on Ethics & Governance of Artificial Intelligence for Health, p 109 and throughout

99 Ethics Guidelines for Trustworthy AI, independent High-Level Expert Group on AI (HLEG), 2019

100 As voiced for example by expert group member Thomas Metzinger: Metzinger, T. (2019). EU

by them all, the recital keeps them firmly away from legal enforcement: '[t]he application of these principles should be translated, when possible, in the design and use of AI models'.

From our point of view, an important consequence of the recitals' on-off, cloudy and selective engagement with the R&D spheres, whilst maintaining that there are strong governing principles in place to keep people safe, is an obfuscation of what the legal regime actually affords to people. The less that law explicitly expresses its ties with the various principled spheres that ground and govern it (e.g. legal and constitutional principles, treaty regimes, and ethical frameworks), the more people are solely dependent on those who do have that knowledge. In light of how legal literacy is already typically low, knowledgeability around legal principles will be even lower. This way law becomes a 'doctor knows best' sphere just like medicine is struggling to no longer be, and just like dominant parts of the AI Industry struggle to maintain.

## 5. Conclusion

This chapter invited readers to engage with the regulation of AI through the critical angle of the philosophical domains of epistemic justice and injustice. Lawmaking can be understood as not only a practice and an instrument of regulatory power but also as one of epistemic power and in/justice. Law adopts, creates or refuses conceptual and ideological understandings of people, regulated subjects, and territories. With knowledge as both a source and an outcome of the process, lawmaking, knowledge-making, and world-making are related dominating activities.

The field of AI, and computational fields before it, are also known for the epistemic injustices that have defined their culture and a part of their output, fuelling the need for regulation. These long-standing problems have come to the fore though many instances of, among others, racist, sexist, casteist, and ableist harm in a world increasingly entangled with computational systems that consolidate existing and historical patterns of injustice and inequity. The industry tends to perpetuate the narrative that these harms are accidental or stem from misuse, and that they come with the rapid advancement of technological complexity – given time, they will go away. Other voices disagree, and the need for a fundamental change in the industry is thoroughly argued for.

guidelines: Ethics washing made in Europe, Tagesspiegel, 8 April 2019 <https://www.tagesspiegel.de/politik/ethics-washing-made-in-europe-5937028.html> see also AlgorithmWatch on the subject: <https://algorithmwatch.org/en/industry-defuses-ethics-guidelines-for-artificial-intelligence/>.

We argued that the EU's regulation of AI provides a crucial moment for the EU law maker to engage with the field of AI in an epistemically-just way. This includes relating the region's responsibility to its historically problematic role as a norm-setter for both regulatory and knowledge-making practices, (and also for law's lack of capacity to deal with systemic injustices) and its dominant position today in the field of technology regulation.

Based on the freshly adopted text of the EU's AI Act, this chapter offered ideas on how to analyse the EU's regulatory effort in terms of its epistemic choices and impacts. With reflections based on illustrations sourced from selected parts of the Act, we hope our chapter provides a rich array of findings to engage with in the imminent life of the Act, offering a specific frame for examining how AI gets do be 'done' and how to protect people from AI's fallout.

In our analyses we explored the Act's overall rejection of grounded, critical understandings of both AI and law. The Act endorses biased, techno-optimist claims of beneficial AI impact and downplays extant harms. The Act places a premium on enabling companies to conduct self-assessments, especially for high-risk applications of AI, and excludes most contemporary harmful applications from its regulated sphere. Exceptions for complying with risk assessments exist for law enforcement with regard to several existing technologies. Critical insights on the failure of legal protections specifically against systemic harm are almost entirely ignored, as individual autonomy is prioritised in line with the current paradigm. Formulated more bluntly, this is the EU making it clear to many groups of people that what they experience does not need novel legal intervention; what they experience is business as usual.

We also explored how the Act creates an obfuscating before-and-after-market narrative of how AI affects the world. This is out of line with an established body of research that reveals AI's intimate interaction with the world in all phases of design and deployment, and ignores the extracted, captured, mined, and sourced participation of workers, data subjects and the environment in the creation of AI to all of their detriment.

The same rejection of contrary evidence can be found in how the Act presents the spheres of R&D, testing and development as spaces of adequate self-regulation that have a 'right to be let alone'.[101] The Act does not specify where the application spheres of scientific research standards and more general ethical guidelines start and stop; it emphasises the non-binding nature of the latter and ignores inadequacies of the former. Its own engagement with the science of AI is hard to follow, and a debunked and highly controversial knowledge-making practice is presented as viable, or possibly safe, save misuse. Furthermore, the urgent need of including affected, excluded, critical, and dif-

101 Pun intended. The 'right to be let alone' stems from the Warren & Brandeis Article of the same name of 1890, widely cited as the origin of individual privacy law as we know it.

ferently-disciplined participants and perspectives in these spheres is only referred to piecemeal and not very strongly; nowhere is it a demand.

Lastly, there is the laudation of European values and European technology governance in an Act that excludes those in critical need of the territory's protection, many from regions who suffer consequences of the territory's industrial and imperial behaviour. The lack of prohibition of the manufacture and export of harmful AI systems goes hand in hand with the negative extraterritorial effects of the law itself. Understanding the Act's epistemic dimensions thus helps to understand how the Act can or cannot cater to other sites and other locations in need of a more socially-just AI practice.

CHAPTER

# XI

# AI Technologies and Discrimination from a Non-Domination Perspective

**Bart van der Sloot,**

**Merel Noorman &**

**Linnet Taylor**[1]

https://doi.org/10.26116/d9dn-yz64

1 Associate, Assistant and Full Professor, Tilburg Institute of Law, Technology, and the Society (TILT), Tilburg Law School (TLS), Tilburg University.

## 1. Introduction

The now famous study of the performance of facial recognition algorithms on a wide range of individual faces by Buolamwini and Gebru neatly illustrated how AI algorithms can lead to intersectional discrimination.[2] They showed that these algorithms performed significantly worse recognising the faces of darker-skinned people than those of lighter-skinned people. Performance dropped even further for darker-skinned women. The study reinforced critiques of the increasing scale of the use of data-driven AI technologies, and reiterates that these systems have an inherent bias that can lead to discrimination against particular groups. However, it also showed that this issue cannot be reduced to just one attribute. Discriminatory AI-based systems can have severe consequences for intersectional groups in society, as they can weaken the political power of these groups in society, and make them disproportionally more vulnerable to interference.[3] This kind of AI-based discrimination also challenges existing legal frameworks, as several scholars have pointed out.[4] This chapter will further explore these challenges and critically analyse the suggested approaches to addressing them, while also outlining what is still missing to effectively protect individuals and groups against the arbitrary use of power by state and non-state actors.

Discrimination is a key concern in the deployment of the recent generations of AI technologies. These data-driven systems are built to classify and categorise people to offer them particular services or make decisions about them. However, when that categorisation is based on features such as race or religion, it becomes morally and legally problematic. At the same time, these new technologies enable categorisations and classifications that can be harmful in new ways that have been underexposed and unaddressed.

Discrimination, in a general sense, refers to treatment that wrongfully disadvantages people based on some salient characteristic such as race, religion, age, ability or gender. It can cause many kinds of individual harm, including physical and psychological, but it can also cause harm to groups when it weakens the power these groups have in exercising their rights and voice in society. In this chapter, we adopt a freedom as non-domination perspective to capture such harms and express them in terms of infringements

2 Buolamwini, J., & Gebru, T. (2018, January). Gender shades: Intersectional accuracy disparities in commercial gender classification. In *Conference on fairness, accountability and transparency* (pp. 77-91). PMLR.

3 Pettit, P. (1996). Freedom as antipower. *Ethics*, *106*(3), 576-604.

4 Wachter, S., Mittelstadt, B., & Russell, C. (2021). Why fairness cannot be automated: Bridging the gap between EU non-discrimination law and AI. *Computer Law & Security Review*, *41*, 105567. Zuiderveen Borgesius, F. J. (2020). Strengthening legal protection against discrimination by algorithms and artificial intelligence. *The International Journal of Human Rights*, *24*(10), 1572-1593.

of freedom.[5] Freedom as non-domination essentially entails that a person's freedom is not restricted arbitrarily by an uncontrolled party in power, such as the government, a private organisation or a fellow citizen. The theory of non-domination focuses not on an actual interference but on the potential for arbitrary interference. It is the lack of restraints, checks and balances, rules on how, when and why power can be used, and the absence of post hoc judicial assessment that is the core focus of freedom of non-domination. This kind of freedom thus requires that power is controlled by those that experience interference by this power. Discrimination can limit freedom in this sense as it constitutes an arbitrary use of power if unjustified and unchecked, and it can weaken the position of particular individuals and groups in regard to their capacity to control those in power.[6]

Under the European legal acquis, discrimination is not illegal per se. The existing legal framework sets the conditions under which people can be treated differently, even if it is based on sensitive features. It is occasionally necessary to do so. Discrimination is only prohibited under particular conditions, which can differ depending on the context. European law bases these prohibitions on a set of criteria that were developed in an age when people, governments and organisations did not have access to the sophisticated data-driven technologies of today. Moreover, these criteria were based on preventing particular concrete and mostly individual harms.[7] As will become clear, the basic tenets that underpin this framework are no longer sufficient to address the problems that emerging AI technologies create.

There is a remarkable feature where the theory of non-domination and the European Convention on Human Rights (ECHR) discrimination law align: they both require the use of power to be rational, necessary, and to have good grounds in order to justify an infringement of rights. Non-domination focuses on clear rules on how, when and why power can be used. This allows citizens to understand how their rights and freedoms may be limited and take it into account when deciding on future courses of action. The arbitrary use of power has no rationality as it is random. Non-discrimination law, in a way, emphasises the same. Governmental agencies can make distinctions and differentiations between groups and people when they exert power. However, these distinctions have to be rational in light of the goal pursued. If law enforcement agencies focus more on men than on women, for example, this is deemed lawful because this distinction is relevant to the goal of fighting crime. After all, men are significantly more prone to

5 Pettit, P. (1996). Freedom as antipower. *Ethics*, *106*(3), 576-604.'

6 In theory, this can be remedied by having representative organisations that can protect the vulnerable. In practice, however, this is can be difficult. The equality bodies established with the intention to perform this function have proven to be ineffective.

7 See on this point: Somek, A. (2011). *Engineering equality: An essay on European anti-discrimination law*. OUP Oxford.

engage in criminal behaviour than women. The random use of power, however, could be considered disproportionate if many people who are not prone to engage in unlawful conduct would be affected. If a police unit were to focus on people with yellow sweaters because the head of the department recently had a negative experience with someone with a sun-coloured hoodie, this would be a prohibited form of discrimination.

The other key aspect of freedom as non-domination is that power should be controlled: there should be checks and balances. Clearly, this raises the question of who sets those checks and balances and who assesses whether the executive adheres to them. If those in power set the rules, the criteria for using power, oversight, and checks and balances would be of little value. This is why non-domination theory emphasises civil participation in public discourse and the right to be heard in concrete decision-making, alongside accountability. This is mirrored in the legal regime, which emphasises the internal correlation between democracy and the rule of law, and it sets out a high number of procedural requirements if citizens' human rights are to be interfered with.[8]

The following section discusses how AI can lead to harmful discrimination, including intersectional and new kinds of discrimination. The third section discusses the legal status quo as regards discrimination law. The fourth section identifies gaps between the technological and legal realities and potential ways to close these gaps. Finally, the fifth section highlights potential paths forward.

## 2. ANTI-DISCRIMINATION PROVISIONS ARE WRITTEN FOR HUMAN DECISION-MAKING

In recent years, there have been numerous examples of how data-driven AI-based systems can lead to discrimination against particular groups of people or individuals, such as the infamous COMPAS system.[9] COMPAS is a risk-prediction system developed by the company Northpointe (now Equivant). It is used by several courts in the US to provide decision support on bail hearings. The system provides a risk score that indicates the predicted probability that the defendant will re-offend based on different kinds of data, including previous convictions and a questionnaire filled out by the defendant. The NGO ProPublica analysed the results of the systems and showed that African Americans were disproportionately given higher risk scores based on comparable track records.[10] There are many

8 Jong, T. D. (2017). Procedurele waarborgen in materiële EVRM-rechten (diss. Leiden). *Deventer: Wolters Kluwer*.

9 Corbett-Davies, S., Pierson, E., Feller, A., & Goel, S. (2016). A computer program used for bail and sentencing decisions was labeled biased against blacks. It's actually not that clear. *Washington Post*, *17*.

10 Larson, J., Mattu, S., Kirchner, L., & Angwin, J. (2016). How we analyzed the COMPAS recidivism algorithm. *ProPublica (5 2016)*, *9*(1), 3-3.

more examples like this, where the use of AI-based systems leads to exclusion or unfair treatment of particular groups of people, including job vacancies being withheld from women, or advertisements for housing being only accessible to certain groups.[11]

The recent examples of discriminatory AI systems highlight some of the features of these systems that can lead to discrimination. It should be noted that discriminatory results are not something exclusive to the use of AI technologies. Earlier algorithmic systems and analogue technologies could also intentionally or unintentionally disadvantage specific groups. For instance, consider the Southern State Parkway's low-built bridges which prevented people reliant on buses for transportation from using the highway and visiting Jones Beach in Long Island in the US.[12] Alternatively, the zoning rules in US cities have exposed communities of colour to more toxic pollutants from industry and traffic than white communities for at least a century if not more.[13] Nevertheless, the kinds of data-driven machine learning technologies that drive the current advancements in AI have the potential to cause harmful discrimination in new ways and at a larger scale because of several characteristics. These characteristics include reliance on inherently biased algorithms, data, proxies and correlations; opaque and complex models; the processing of large volumes of heterogenous data; and unpredictable interactions with real-world environments.[14]

Bias in algorithms and data is the most common source of discriminatory treatment, and it is an inevitable part of these systems. During their development choices have to be made, for instance, about what data should be used, what these data represent and what kind of models are best suited for the given data. Every choice introduces bias in the sense that bias necessarily results from the process of sampling and shaping a dataset for analysis. Models and data are abstractions of real-world phenomena of interest; each abstraction highlights particular features of phenomena at the cost of others. In other words, there is 'no free lunch' in model optimisation.[15] To have a Machine Learning (ML) system learn a particular model, choices will have to be made about certain trade-offs, such as whether to minimise false positives or false negatives. These decisions always accompany the development of AI systems, and they always come at a price. The choice of one comes at the cost of the other. Thus, choosing one particular

11 Noble, S. U. (2018). Algorithms of oppression: How search engines reinforce racism. In *Algorithms of oppression*. New York university press. O'neil, C. (2017). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.

12 Winner, L. (2017). Do artifacts have politics?. In *Computer ethics* (pp. 177-192). Routledge.

13 Jbaily, A., Zhou, X., Liu, J., Lee, T. H., Kamareddine, L., Verguet, S., & Dominici, F. (2022). Air pollution exposure disparities across US population and income groups. *Nature*, *601*(7892), 228-233.

14 Gerards, J., & Xenidis, R. (2021). *Algorithmic discrimination in Europe: Challenges and opportunities for gender equality and non-discrimination law*. European Commission..

15 Wolpert, D. H., & Macready, W. G. (1997). No free lunch theorems for optimization. *IEEE transactions on evolutionary computation*, *1*(1), 67-82.

fairness metric to measure whether recidivism risk scores are unbiased may come at the cost of other conceptions of an unbiased distribution. Technologists and statisticians have offered various definitions of fairness to assess the distribution of outcomes across different groups, where one might for example focus on the equal distribution across members of different groups and others might focus on maximum accuracy for every individual, regardless of group membership. The different definitions, however, are not always compatible, so choices have to be made.[16] Similarly, data always have a bias. They reflect the interests, world views and priorities of those who collect, select and categorise them. Data represent measurable aspects of the world and what they represent is the result of a choice: a choice to capture or measure these aspects in a particular way. This process introduces biases. Nevertheless, not all biases are alike, and some biases are more problematic than others, as they can lead to discriminatory and harmful outcomes, intentionally or unintentionally.

As this would all suggest, bias can occur at various stages of the lifecycle of an AI model. Collected data can reflect existing biases or over- or underrepresent particular relevant features in the world. Bias can result from a prejudiced choice of labels for data or if the performance metrics of a model are ill-suited to how the model be will applied. It can also occur if the outcomes of a system are interpreted in a particular light or are used to target particular populations, creating a disproportionate advantage or disadvantage for other groups.[17]

The biases that affect the design, development and use of AI systems are inextricably linked to biases that occur in sociocultural and institutional contexts. Computer developers and data scientists are not free from bias in their understanding of the world. Their backgrounds and interests colour their framing of particular problems and their interpretation of successful outcomes. Moreover, they tend to be primarily trained to develop technical skills rather than legal, social and ethical expertise. This social bias on the individual level is connected to social bias on the organisational level. The lack of diversity in AI teams hampers the correction of problematic biases in the development and application of AI because of the tendency of human beings to for example search for or interpret information in a way that confirms personal preconceptions.[18] Another kind of bias results from the pervasive techno-optimism in many companies and organisations. The strong belief in the ability of technology to solve all sorts of problems

16 Chouldechova, A. (2017). Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big data*, *5*(2), 153-163.

17 Suresh, H., & Guttag, J. (2021, October). A framework for understanding sources of harm throughout the machine learning life cycle. In *Proceedings of the 1st ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization* (pp. 1-9).

18 Kuhlman, C., Jackson, L., & Chunara, R. (2020). No computation without representation: Avoiding data and algorithm biases through diversity. *arXiv preprint arXiv:2002.11836*.

prevents further reflection on the appropriateness or suitability of AI solutions to particular problems.

These different forms of bias show that it is not just data that are biased. Bias can also be the result of choices made in algorithm development, the training of models and the implementation and use of technology. Even if it were possible to develop unbiased datasets, discriminatory outcomes may still occur because of the choices made while designing algorithms and optimising models.

The different kinds of biases also illustrate that AI-based discrimination can be the consequence of blind spots, implicit assumptions and intentional decisions. Implicit bias can occur, for instance, in the design of an AI system, where developers assume the average user has a certain level of knowledge and capacities. This may exclude the visually impaired from using the system effectively. However, data-driven learning systems can also have biases if the system is built to find correlations that can serve as proxies for ethnicity, gender or religion. In a study of Google Adsense, Sweeney demonstrated how names that were more prevalent in African-American groups generated more negatively framed targeted ads than names that were more popular amongst white Americans. [19] This discrimination in targeted advertisements also occurs on other platforms. To further fine-tune the targeting of ads for individual users, companies like Google and Facebook analyse data to infer affinity groups that can serve as proxies for ethnicity, sexual preference and gender. A person who regularly visits feminist sites can be assigned to the affinity group 'woman', for example. This kind of targeting leads to certain groups being excluded from seeing certain advertisements, for instance for financial services, housing or job vacancies. This can compound existing societal and political power asymmetries, and it places certain groups at considerable disadvantage.[20]

Data-driven AI systems do not just have bias, they can also reinforce and exacerbate existing biases and disparate treatment of particular groups. This can happen, for example, when human operators either trust the outcomes of systems too much or too little. For example, a judge under time pressure may rely too much on the risk-of-recidivism score generated by a decision-support system to decide on bail for a defendant. They may fail to critically reflect on or interrogate the outcomes of the systems, and follow the results blindly. If these outcomes disadvantage particular groups, this automation bias can widen and strengthen the discriminatory treatment of these groups. This can even be exacerbated if these systems are used to scale up decision-making processes and increase their efficiency.

19 Sweeney, L. (2013). Discrimination in online ad delivery. *Communications of the ACM*, *56*(5), 44-54.

20 Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A., & Rieke, A. (2019). Discrimination through optimization: How Facebook's Ad delivery can lead to biased outcomes. *Proceedings of the ACM on human-computer interaction*, *3*(CSCW), 1-30.

Moreover, the ability of these systems to analyse large volumes of different kinds of data and find complex correlations and patterns increases the chance of intersectional discrimination occurring.[21] Today's AI systems can combine multiple characteristics to automatically derive very specific categories and profiles to offer services or assess risks, for example. In this way, they can also create new kinds of discrimination, as the categories that these systems identify might not necessarily be recognisable as meaningful human characteristics. As a result, individuals might be unfairly treated based on a membership of some newly-emerged salient group centred on certain traits, features, preferences or behaviours they are unaware of.[22] Developers might not always be able to anticipate these kinds of discrimination, as such groups may only emerge if the system is applied in practice. One example would be a situation where culture determines seemingly unrelated preferences, for instance where particular online purchasing habits turn out to correlate with a protected characteristic like ethnicity. If this information is used to target ads about jobs or housing, this could create significant disadvantages for the particular group in almost invisible ways.[23] These systems make it hard to detect discrimination because the complexity of the models used and the fact that they run quietly in the background of other processes obstruct a clear view of the logic that produced a particular result, or even the results themselves.[24] This increases the risk of the arbitrary use of power on particular groups of people.

Thus, although discrimination stemming from current day AI technologies is not an entirely new phenomenon, several of their characterising features increase the risk of discrimination and even challenge existing ways of addressing these risks. In the following section, we will examine how the current European legal framework falls short in addressing these risks.

## 3. Anti-Discrimination: A Rationality Discourse

The modern human rights discourse arose from the ashes of the Second World War.[25] The United Nations adopted the Universal Declaration on Human Rights in 1948, the International Covenant on Civil and Political Rights in 1966, and the Council of Europe

21 Crenshaw, K. W. (2013). Mapping the margins: Intersectionality, identity politics, and violence against women of color. In *The public nature of private violence* (pp. 93-118). Routledge.

22 Wachter, S., & Mittelstadt, B. (2019). A right to reasonable inferences: re-thinking data protection law in the age of big data and AI. *Colum. Bus. L. Rev.*, 494.

23 Cossette-Lefebvre, H., & Maclure, J. (2023). AI's fairness problem: understanding wrongful discrimination in the context of automated decision-making. *AI and Ethics*, *3*(4), 1255-1269.

24 Wachter, S. (2020). Affinity profiling and discrimination by association in online behavioral advertising. *Berkeley Technology Law Journal*, *35*(2), 367-430.

25 Some thoughts in this section are also contained in: van der Sloot, B. (2024). *Regulating the Synthetic Society: Generative AI, Legal Questions, and Societal Challenges* (p. 296). Bloomsbury Academic.

the European Convention on Human Rights in 1950. A non-discrimination clause featured prominently in each of them, the atrocities committed by fascist regimes that targeted specific groups based on race, sexual orientation, mental capacity and religion still fresh in mind. Recalling 'the horrible triumph of discrimination on grounds of race or political opinions', [26] the authors of the ECHR specified in Article 14:

> The enjoyment of the rights and freedoms set forth in this Convention shall be secured without discrimination on any ground such as sex, race, colour, language, religion, political or other opinion, national or social origin, association with a national minority, property, birth or other status.

To understand how this provision applies to the AI context, the background of this provision and the way in which it is interpreted by the European Court of Human Rights is discussed in this section. Three general points will become clear in doing so. First, the anti-discrimination provision addresses concrete harms, mostly on an individual level. This means that systemic problems and abstract harms are difficult to address under the current regulatory regime. Second, the provision is written for human decision-making. It specifies grounds that historically humans unfairly discriminate. It is grounded in the limited human capacity to understand how decisions come about and what their effects are. Third, rather than the question of discrimination, the central question under Article 14 ECHR is relevance. Discrimination and differentiation are not only accepted, they are legally required: the question the Court will ask when dealing with claims is thus not so much 'is this discrimination or not?' but rather 'is the discrimination between groups relevant or not?'. This section mainly draws from the European Court of Human Rights' (ECtHR or the 'Court') jurisprudence, with several more concrete examples taken from the Dutch legal context by way of illustration.

### 3.1 Anti-Discrimination Provisions Address Concrete and Demonstrable Individual Harms

To understand the right to non-discrimination under the Convention, it is important to start with a basic decision the Court has made in dealing with submissions under the ECHR. In principle, applications will only be declared admissible if citizens can successfully demonstrate that they have suffered harm to one or more interests that fall under the scope of the ECHR. This harm must be substantial.[27] For example, although unlaw-

26 Robertson, A. H. (Ed.). (1985). *Collected Edition of the'travaux Préparatoires' of the European Convention on Human Rights*. Martinus Nijhoff.

27 Article 35 § 3 sub (b) ECHR. Meyer-Ladewig, J. (2012). The Principle of De Minimis Non Curat

fully processing a citizen's name and address is *stricto sensu* a violation of Article 8 ECHR (the right to privacy), a submission on this point would be rejected by the Court because the violation of the law only resulted in minimal harm. In addition, groups are not allowed to submit a claim under the Convention, for example to protect their group interest, though each individual member can. This means that, for example, if Roma are discriminated against, the Roma community cannot submit a claim *qualitate qua*, although each individual member can.[28] These members must then each be able to demonstrate that they have individually suffered harm from the discriminatory policy or practice.

Both points have important repercussions for the AI context. Inter alia, this means it is difficult to address systemic racism as such.[29] In principle, the Court only assess submissions on a case-by-case basis, assessing the harm done to a particular person in a particular context. This also means that the Court will only issue decisions based on what needs to be done in order to remedy a harm in a particular case. This means that the underlying system can remain in place. For example, generally, it will not assess the extent to which there is institutional racism and, if so, what needs to be done to revise the organisational structure. Instead, it assesses whether a person was discriminated against in a particular case and, if so, what should be done to remedy that violation, such as undoing the discriminatory effects and/or providing monetary damages.[30] This also means it is up to an individual to recognise harm and recognise that it is the result of a discriminatory practice. However, this may be far from clear, given that lived experiences often do not align with official statistics and seeing as courts will prioritise the latter as grounded truth. This problem is deepened in the context of AI, among others because of how the AI system operates and because the data on which it runs is generally not communicated to individuals.[31] Finally, it means that most anti-discrimination cases are

Praetor in the Protection System of the European Convention on Human Rights. *Const. L. Rev.*, 5, 127.

28 Even of a Member State to the Council of Europe allows for class actions, this does not mean that class actions will be declared admissible by the European Court of Human Rights. This has been analysed in detail in: Van der Sloot, B. (2017). Do groups have a right to protect their group interest in privacy and should they? Peeling the onion of rights and interests protected under article 8 ECHR. Taylor, L., Floridi, L., & Van der Sloot, B. (Eds.). (2016). *Group privacy: New challenges of data technologies* (Vol. 126). Springer.

29 <https://publicaties.mensenrechten.nl/file/a261a614-6d6e-4d1d-be8c-0c88fd43b954.pdf>

30 There are experiments with pilot judgements that try to remedy this flaw in the Convention mechanism. 'Pilot-Judgment Procedure.' <https://www.echr.coe.int/documents/d/echr/pilot_judgment_procedure_eng#:~:text=The%20central%20idea%20behind%20the,an%20individual%20basis%20in%20Strasbourg.>. Though there is a deterrent effect of judgement, this effect has been watered down significantly due to changes made to the Convention mechanism over time: Van der Sloot, B. (2017). *Privacy as virtue*. Intersentia.

31 See on this point, to the contrary, the relevant case law of the ECtHR under Article 8 ECHR on the

matters of 'David against Goliath', so to speak. Most individuals do not have the time, expertise or money to sit out long legal battles against big companies or governmental agencies that use AI.

A second point that relates to the first is that the Court is very focused on the concrete harm and negative consequences of discriminatory practices. Accordingly, in principle, harm should materialise or be directly foreseeable in the eyes of the Court for it to issue a judgment. In principle, citizens cannot submit *a priori* or hypothetical claims. This ties in with a specific point about the anti-discrimination clause in the ECHR, and it is very focussed on taken decisions or adopted policies and the effects thereof. In general, the Court differentiates between direct discrimination (decisions taken based on one of the factors listed in Article 14 ECHR) and indirect discrimination (decisions not taken based on of one of those factors that nevertheless have a substantial negative impact on groups with a certain racial, political, sexual, etc. background). There are few procedural safeguards the Court has laid down, as it looks more to questions of material than procedural justice.[32]

This has clear effects for the applicability of non-discrimination law in the context of AI. Most codes and frameworks on ethical AI focus on the process leading up to a decision. For example, they focus on which data are collected, how they are categorised, which algorithm is deployed, how correlations are interpreted and how diverse the group operating the AI system is. This part of the process, however, does not fall under the scope of non-discrimination law. Clearly, humans have limited experiences, access to biased data, and use imperfect decision trees when making decisions or policies. However, these problems are seldom recognised, as they are opaque and almost impossible to scrutinise. With AI, biases may become apparent, and it may become clear that, just like most humans, AI is always implicitly biased, as it is also based on limited and biased data. Although these biases are often problematised within the context of AI, Courts seldom scrutinise the biases that influence policies and decisions.

This means that large parts of the process are left unregulated and that harm must already have materialised before a citizen can address a violation of law.[33] Consequently,

quality of law and *in abstracto* claims, e.g. ECtHR, Zakharov v. Russia, application no. 47143/06, 04 December 2015. ECtHR, Big Borther Watch and others v. the United Kingdom, application nos. 58170/13, 62322/14 and 24960/15, 25 May 2021.

32 See again, to the contrary, the relevant case law of the ECtHR under other provisions, such as the four qualified rights, where it has found procedural requirements to be implicit in the substantive provisions. See for example Arnardóttir, O. M. (2017). The 'procedural turn' under the European Convention on Human Rights and presumptions of Convention compliance. *International Journal of Constitutional Law*, *15*(1), 9-35. Cumper, P., & Lewis, T. (2019). Blanket bans, subsidiarity, and the procedural turn of the European Court of Human Rights. *International & Comparative Law Quarterly*, *68*(3), 611-638.

33 See for example ECmHR, Tauira and others v. France, application no. 28204/95, 04 December 1995.

the underlying bias in systems cannot be addressed as such. It also means that some parts of AI systems fall outside the scope of the Convention altogether, for instance when decisions or system effects are not directly relatable to concrete individuals. This may be the case, for example, because AI systems are used to structure internal processes, take general decisions that affect the population at large or have positive effects on specific groups. One example of the initial part of structuring such internal processes could be predictive policing systems that suggest that police units should monitor certain areas. Although the system may be biased, having police officers walk through one's street more often than in other neighbourhoods is not legally considered a harm. If the police officers subsequently arrest someone, the question that the Court will assess is whether the police had good grounds to arrest that person. The Court also assesses whether the police officer had good grounds to walk through that specific street at that specific time. This is considered a political or policy matter, not a legal one. An example of taking general decisions that affect the population at large is if AI is used to determine what bridges should have priority in terms of repairs or replacement. Such a system may be biased in favour of repairing bridges in affluent neighbourhoods. However, under the current legal status quo, this bias will may raise a moral or political issue, but not a legal one. One example of the third, having positive effects on a specific group, is when insurers treat everyone equally, but it lowers the monthly payments for a specific group based on insights from its AI system. Since this group does not suffer from negative consequences and the other group, of say 90% of the customers, is not homogeneous in terms of race, political or sexual preference and so forth, these effects of AI systems will most likely fall outside the scope of the non-discrimination regime.[34]

### 3.2 Anti-Discrimination Provision is Written for Human Decision-Making

The second relevant observation is that anti-discrimination law is written for human decision-making. This is evident, first, from the limited number of grounds engrained in Article 14 ECHR. This list is the result of the ambition to protect citizens against the kinds of discrimination that individuals have historically suffered from. Although there is a 'catch-all' provision that refers to 'other status', the European Court of Human Rights has only been prepared to accept grounds under this category that are related to the grounds explicitly mentioned. These include age, gender identity, sexual orientation, health and disability, parental and marital status and immigration status.[35] Although it

34 Under EU law, there has been a relevant case on this point: CJEU, Association belge des Consommateurs Test-Achats ASBL, Yann van Vugt, Charles Basselier v Conseil des ministres, In Case C236/09, 1 March 2011, ECLI:EU:C:2011:100.

35 <https://www.echr.coe.int/documents/d/echr/Handbook_non_discri_law_ENG_for_AZE>.

is true that many of these grounds are aspects that people have limited influence on, this may not hold true for several other grounds, such as political beliefs. These grounds are listed because they have been abused for discriminatory policies historically.

This has led to complicated questions when explicitly discriminatory policies are based on other factors. These factors might include the type of smartphone a person has, the colour of their hair or their shoe size. For example, if border police could only pick out people with shoe size 41, this may be difficult to address under the current regulatory regime. However, it may be argued that this criterion has an indirect impact on gender or ethnic background. In any case, it is clear that the grounds listed in Article 14 ECHR are difficult to apply to the context of AI. AI systems can have discriminatory effects on the groups mentioned in the European Convention on Human Rights, when fed with historically biased datasets and operated by biased human operators. Yet AI systems are not implicitly inclined to base decisions on ethnicity, sexual orientation or political beliefs. Rather, they may take decisions on any criterion that has been deemed statistically relevant. This can raise difficult questions, and is already causing harm. While humans develop discriminatory policies based on ethnicity, sexual orientation, religion and so forth, this is not necessarily how AI systems 'inherently' discriminate. Accordingly, it is questionable whether these aforementioned grounds still suffice in an environment where many decisions and policies are wholly or partially based on algorithmic processes.

Anti-discrimination law is not only developed to prevent the type of discrimination that has historically taken place. It is also based on the belief that the number of factors a policy or decision is based on is relatively small, and that these factors are relatively stable. This assumption results from human limitations in terms of mental capacity and the practical investments in time, energy and resources required to change and update the criteria relevant to policy and decision-making processes. The anti-discrimination regime is based on limited human capacity in another sense as well. In terms of direct discrimination, it only considers grounds made explicit. It does not consider implicit discriminatory grounds or grounds people are unaware of. Although we know that many people are subconsciously racist, for example, anti-discrimination law generally does not assess the extent to which a police officer's decision to stop and search a Black man's vehicle was subconsciously motivated by racial motives. Obviously, when a police officer or a police unit discriminates against a Black man subconsciously or otherwise, this can and should be treated under the doctrine of indirect discrimination. This doctrine evidently focuses on the effects, rather than implicit assumptions of a decision. However, there are important practical limits to assessing what practical effects policies and decisions have on various groups in society. This is both because of limited data about the effects of decisions and policies and because of the limited resources for making such assessments.

AI poses new challenges to the standard approach to anti-discrimination law. In principle, policies can be created and decisions can be taken not based on five relevant factors, but on 5,000. Clearly, if one or more of the five factors is discriminatory in nature, this may raise a red flag. However, the larger the number of factors and the smaller the relative weight of the discriminatory ground, the less evident it is. In addition, in AI systems, the factors that are used to make decisions may change incredibly fast:

> in the big data era, groups are increasingly fluid, not only through their changing membership, but also because of the changing criteria for the group itself. A group, the criteria for grouping people and the membership of a group might change in a split second. The purpose for which the group is designed may also change from day to day to adapt to new insights gained from data analytics, and groups may be formed and dissolved through the push of a button. [36]

In today's data-driven environment, it is possible to assess the effects of policies and decisions on groups in detail. It will become clear that any data point is indirectly correlated to a discrimination ground. For example, a certain shampoo might be bought for 80% by women, 18% by men and 2% by non-binary people, for 67% by Black people and for 22% by white people, for 30% by Christians, 31% by Muslims and for 33% by atheists, etc. The question is to what extent all these indirect correlations should be considered. Every policy will also have different effects on different groups. A policy to grant tax benefits to people living in rural areas, for example, may have the effect that white males are the main beneficiary because immigrants and single women tend to relocate to municipal areas. A city's policy to subsidise cultural institutes may benefit women more than men because the latter, on average, frequent museums, concert halls and theatres less than women do. All policies will have indirect discriminatory effects on all groups. Again, the more data that is available about the effects of policies and decisions on different groups, the more urgent the question about where to draw the line is.[37]

36 Taylor, L., Floridi, L., & Van der Sloot, B. (Eds.). (2016). *Group privacy: New challenges of data technologies* (Vol. 126). Springer, p. 284.

37 Discrimination law addresses indirect discrimination, which concerns the negative effect of policies (intended or unintended) on marginalised groups in society, even when no discriminatory grounds are used. See for example Tobler, C. (2005). *Indirect discrimination: a case study into the development of the legal concept of indirect discrimination under EC law* (Vol. 10). Intersentia. Etinski, R. (2013). Indirect Discrimination in the Case-Law of the European Court of Human Rights. *Zbornik Radova*, 47, 57. Besson, S. (2012). Evolutions in non-discrimination law within the ECHR and the ESC systems: It takes two to tango in the council of Europe. *The American Journal of Comparative Law*, *60*(1), 147-180. Zuiderveen Borgesius, F. J. (2020). Strengthening legal protection against discrimination by algorithms and artificial intelligence. *The International Journal of Human Rights*, *24*(10), 1572-1593.

This is connected to the fact that Article 14 ECHR is not primarily written as a subjective right but as an objective right. This means it is an obligation on parties not to discriminate.[38] Although the Court has acknowledged that citizens can directly rely on Article 14 ECHR, it is only considered an ancillary right. Unlike all other rights under the Convention that can be invoked by citizens independently, Article 14 cannot. Accordingly, citizens cannot complain about discrimination as such. They can only do so when one of their other rights has been interfered with on discriminatory grounds. For example, when only the homes of certain groups in society are searched, this may lead to a violation of Article 8 and 14 ECHR. It is not considered discrimination as such, which is problematic under the Convention. It is only problematic when the state violates human rights in a discriminatory way.[39]

Evidently, this principle already has limited practical meaning for non-AI based forms of discrimination, as discrimination done by humans varies in degrees of tangible, concrete effects on a person's freedoms and rights. This lack of tangibility, in addition to a lack of traceability, is even more pronounced when it comes to AI systems. Accordingly, when an AI system shows Black people pictures of hungry children in a doctor's waiting room because it has found a 'relevant' correlation, while it shows pictures of Nobel prize winners to white people, while clearly problematic, it is not easy to see under the scope of which provision of the Convention this practice could be addressed. It also means that discriminatory issues will not fall under the regulatory regime per se because autonomy is not directly covered by the Convention, even though it may be a relevant aspect in relation to the right to privacy. To give a final example, the aggregate or cumulative effects of AI systems may be difficult to handle under the non-discrimination regime. Although the effect of a singular decision or policy may not be seen as significant enough to bring it under the scope of the Convention, the cumulative effects of AI systems that negatively affect disenfranchised groups may have significant repercussions (the reverse Matthew effect). However, each individual decision or policy may in and of itself be considered insufficiently harmful to bring it under the scope of the ECHR.

38 Gerards, J. H., Heringa, A. W., & Janssen, H. L. (2005). *Genetic discrimination and genetic privacy in a comparative perspective* (Vol. 51). Intersentia. Gerards, J. (Ed.). (2023). *Fundamental Rights: The European and International Dimension*. Cambridge University Press. Gerards, J. (2013). The discrimination grounds of article 14 of the European convention on Human Rights. *Human Rights Law Review*, *13*(1), 99-124.

39 <https://www.echr.coe.int/documents/d/echr/Guide_Art_14_Art_1_Protocol_12_ENG>.

### 3.3 The Central Question under Article 14 ECHR is Relevance

The third observation about the applicability of the anti-discrimination provision in the ECHR to AI-mediated human action is that human rights instruments do not principally attempt to abolish discrimination. Instead, the central rationale is the abolition of the arbitrary use of power by the state. Perhaps counter-intuitively, the human rights instruments' main concern is laying down principles concerning the rule of law. States can and need to enter a citizens' home, set restrictions on freedom of speech and limit the freedom of citizens as in general. There are situations recognised under the Convention when forced labour, the death penalty or imprisonment without trial can be deemed legitimate. What the Convention requires is that if states interfere with human rights, they do so in a non-arbitrary way. This underlying rationale should be seen as a response to totalitarian regimes where the executive power was able to operate without any legal boundary so it could use power at will. The human rights discourse should be seen as providing safeguards against this sort of situation from occurring again, in particular by requiring a good ground when the executive uses its power.[40]

This also holds true for the non-discrimination provision specifically. The Convention does not prohibit states from discriminating against Black, homosexual, and/or Jewish people or any other marginalised group. It prohibits states from doing so without good grounds. Accordingly, the Court, without interfering with any of the substantive provisions of the Convention and when a potential violation of Article 14 ECHR arises, assesses whether there were good grounds for doing so. It prohibits parties from discriminating and mandates discrimination. Precisely because human rights cannot be violated at will, states can only limit the rights of individuals or groups with respect to the good grounds the states have. This means the police cannot enter unmarried women's homes purely because they hate unmarried women. They have to demonstrate that there are good reasons to enter the home of a specific individual or group. If the police only enter the homes of Jewish people because they have received reliable intel from the Jewish community in a city that several attacks are planned, this may be deemed legitimate by the Court. If the police have information that a man between the ages of twenty and thirty is going to commit an attack at a certain train station, not only can the police use this information as relevant criteria, it could legally be argued that they should. If the police were to do body cavity checks on women over sixty in response to this information, this might be deemed a human rights violation because there were no good grounds for doing so. Following the idea that the state should keep human rights interferences to a minimum, they should operate in the most effective way: they should discriminate.

40 Van der Sloot, B. (2023), Editorial, EDPL 2023-2.

An ethnic profiling case involving the Dutch border police further illustrates the complicated and controversial nature of this provision. This case was initially deemed legitimate by the court because ethnicity was a relevant criterion with respect to the aim of preventing illegal stays in the Netherlands.

The State pointed out that Mobile Surveillance Security (MTV) controls may not have the same effect as border controls at the internal borders. For this reason, they are limited in number, frequency and scope in Article 4.17a paragraph 3 to 5 of the Aliens Decree 2000. For each flight, only part of the passengers may be checked, for each train only part of the train may be searched and in no more than four compartments, and on the road or waterway only part of the passing vehicles or ships may be stopped. This precludes the alternative of blanket checks mentioned by Amnesty International et al. Purely random checks would, so the State argued during the session, greatly reduce the effectiveness of the MTV because action would not be sufficiently targeted. Given the nature and the objective of the MTV – fighting 'illegal' stay in the Netherlands – any alternative for targeted selection decisions has therefore not appeared.[41]

Note that it is difficult for the courts and targeted individuals to dispute these claims if data on these experiments are not made public.

The Court of Appeal arrived at a different conclusion. It did not do so because race or ethnicity cannot be used as a relevant criterion by the state. Instead, it argued that race and ethnicity were used as an exclusive or decisive criterion for which there was no objective or reasonable justification. As such it was unlawful.[42] Moreover, it noted that this kind of discrimination – based on race – enacted through the policy and actions of the Dutch Border police had 'a negative effect on society as a whole. Dutch people with a skin colour other than white can as [a] result not feel accepted and feel like second-class citizens'. This justifies the application of a strict assessment, according to the Court. By noting this, it underlines that discrimination is not just about individual harm, but that the lived experience of individuals and groups affected by discrimination has real societal consequences that should be taken into consideration.[43] This shows a different approach to the dominant individualistic conception of discrimination in the ECtHR's case law. Moreover, it echoes a freedom as non-domination perspective on harmful discrimination by parties in power, in that curtailing harmful discrimination is as much about restricting the arbitrary use of power by these parties as it is about ensuring that everyone affected is equally positioned to control that power. Feeling like a second-class citizen can lead to further marginalisation and can negatively affect the relationship

41 ECLI:NL:RBDHA:2021:10283, Rechtbank Den Haag, C-09-589067-HA ZA 20-235.

42 ECLI:NL:GHDHA:2023:173, Gerechtshof Den Haag, C-09-589067-HA ZA 20-235.

43 See also Opinion 4/2020 on the European Commission's White Paper on Artificial Intelligence—A European Approach to Excellence and Trust. <https://edps.europa.eu/sites/edp/files/publication/20-06-19_opinion_ai_white_paper_en.pdf>.

between these citizens and the state.[44] A lack of access to relevant information on the chosen method, such as its effectiveness, further weakens the position of these citizens.

The Court's decision also elucidates another important aspect: a court will not only assess whether the goal as such is legitimate and whether discriminatory grounds or effects can be deemed necessary in light of that goal, but also whether there are less intrusive means for doing so. This approach is used when courts assess discriminatory policies of private sector organisations as well: a supermarket may not deny job applications from Black people because it believes that they tend to be inherently lazy, or prevent women from attaining higher positions because it holds that they are by definition hysterical and unfit to lead. However, a gentleman's club may base its membership admission policy on gender, and a Christian school may assess the extent to which prospective teachers endorse Christian values.[45] A school that bases its admission policy for students on race because it has found a correlation with early school dropout should legally rather use another factor that also correlates with and has a predictive value for school dropout, if that is something they wish to target.

This means three things for the AI context. First, as the core of anti-discrimination law is to assess whether there were good reasons to discriminate, black box systems are not acceptable in general. If an organisation that deploys an AI system cannot explain why it was necessary and reasonable to discriminate,[46] courts will find this to be a violation. The burden of proof is on the claimant to prove that they have suffered harm from a discriminatory policy or decision. However, the burden of proof is thereafter on the organisation that makes the decision or policy to demonstrate that such differentiation was objectively reasonable. Second, non-black box AI may make it easier for organisations to demonstrate an objectively reasonable ground because it can point to a correlation between one data point (a proxy) and the data point relevant to the decision or policy goal. Third, it may also mean that it should be able to find alternative proxies that are not discriminatory, or are less so, more easily. Here, courts may need to lay down a standard because finding alternative proxies may come at the price of effectiveness.

Suppose the police in City A wants to predict crime. Clearly, it has no data on who will commit crimes so it uses proxies. It has several data points that may assist, which have different predictive values. The highest predictor is the combination of age, gender and ethnicity, as Black boys and men between the age of 15–35 are convicted for 80% of all recorded crimes in City A. The police can also use location-based predictions, as zip-

44 Alston, P. (2018). The human rights implications of extreme inequality. *United Nations, General Assembly, Report of the Special Rapporteur on extreme poverty and human rights, A/HRC/29/31 (27 May 2015), NYU School of Law, Public Law Research Paper*, (18-06).

45 See for example <https://oordelen.mensenrechten.nl/oordeel/2021-59>.

46 There is a positive obligation for Member States under the ECHR to ensure that both private and public sector organisations adhere to the ECHR and the jurisprudence of the ECtHR.

codes have a 70% predictive value for where crimes will take place. However, zip-codes have an indirect correlation with ethnicity. If police activities are based on zip-codes, this would mean that 80% of all people subjected to surveillance activities would be Black, regardless of gender. Alternatively, the policy could be based on time slots, with most crime taking place between 8:00 p.m. and 4:00 am. Yet, this has a predictive value of 60% only and would still have an indirect effect on Black people because they are overrepresented in the group of people with a night-time job, as 65% of this group is Black. In short, it is currently unclear which data point should be chosen – if the system should be used at all – although the courts would presumably allow for a margin of appreciation. The questions that remain are, first, what amounts to indirect discrimination and what does not and second, if the Court establishes indirect discrimination, whether it is deemed proportionate to the goal pursued. Both legal tests are binary. This means there is indirect discrimination or there is not, and an interference is proportionate or it is not, while reality reveals sliding scales.

This leads to a final point: anti-discrimination law is inherently conservative and may reinforce social inequalities. States may explicitly base their policies on discriminatory grounds. Just as a Christian school may found its organisation and policies on its Christian faith, a nation may base its legal system on, for example, Christian values. Accordingly, behaviour that may conflict with Christian teaching may be prohibited or treated differently. For example, many European states still grant a special status and attach legal or financial privileges to married couples over non-married ones. They prohibit marriages between three or more people, and many countries do not recognise same-sex marriages. Although clearly discriminatory, this is allowed by the ECtHR as it deems it legitimate and even necessary for states to create a moral community. In general, this can be seen as the tension between the rule of law and democracy. If 90% of a country's population is deeply appalled by same sex marriages, following the principle of democracy, the legislator cannot and should not ignore that sentiment. Accordingly, one of the officially recognised grounds on which countries can legitimately curtail human rights is the protection of the health and morals of the community. As mentioned, policies can differentiate between marital status. This was made clear when two cohabiting sisters complained of a difference in inheritance tax. Member states are allowed to favour legally recognised forms of cohabitation over others.

> As with marriage, the Grand Chamber considers that the legal consequences of civil partnership under the 2004 Act, which couples expressly and deliberately decide to incur, set these types of relationship apart from other forms of cohabitation. Rather than the length or the supportive nature of the relationship, what is determinative is the existence of a public undertaking, carrying with it a body of rights and obligations of a contractual nature. Just as there can be no analogy between married and

> Civil Partnership Act couples, on the one hand, and heterosexual or homosexual couples who choose to live together but not to become husband and wife or civil partners, on the other hand ... the absence of such a legally binding agreement between the applicants renders their relationship of cohabitation, despite its long duration, fundamentally different to that of a married or civil partnership couple. This view is unaffected by the fact that (...) Member States have adopted a variety of different rules of succession as between survivors of a marriage, civil partnership and those in a close family relationship and have similarly adopted different policies as regards the grant of inheritance-tax exemptions to the various categories of survivor; States, in principle, remaining free to devise different rules in the field of taxation policy. In conclusion, therefore, the Grand Chamber considers that the applicants, as cohabiting sisters, cannot be compared for the purposes of Article 14 to a married or Civil Partnership Act couple.[47]

Accordingly, discrimination on the basis of marital status is deemed a legitimate aim in itself and may be an official policy goal. In the famous Marckx case, regarding differentiation in the national law between the rights of legitimate and illegitimate children to inherit, the government referred to the 'traditional family' and maintained 'that the law aims at ensuring that family's full development and is thereby founded on objective and reasonable grounds relating to morals and public order'. The court, although denouncing any form of discrimination, accepted 'that support and encouragement of the traditional family is in itself legitimate or even praiseworthy'.[48]

The court does override a national legislator at times when it discriminates, but it is very hesitant to do so. When considering these kinds of cases, the ECtHR looks at the 'European consensus' or even the 'international consensus' on a certain point. Accordingly, if most countries in Europe prohibit same sex marriages, single countries are allowed to do so as well. If various countries have adopted a different approach to the issue, countries are allowed a 'margin of appreciation'. Even if only a small number of countries would prohibit same-sex marriages, a country would still be allowed to do so because the Court has stressed that states should be granted a wide margin of appreciation especially when moral considerations are at stake. This also applies to other sensitive moral issues, such as in vitro fertilisation (IVF).

Since the use of in vitro fertilisation treatment continues to give rise to sensitive moral and ethical issues against a background of fast-moving medical and scientific developments, and since the questions raised by the present case touch on areas where there is not yet clear common ground among the member States, the Court considers that the

47 ECtHR, Burden v. the UK, application no. 13378/05, 29 April 2008, § 65.
48 ECtHR, Marckx v. Belgium, application no. 6833/74, 13 June 1979, § 40.

margin of appreciation to be afforded to the respondent State must be a wide one. The State's margin in principle extends both to its decision to intervene in the area and, once having intervened, to the detailed rules it lays down in order to achieve a balance between the competing public and private interests.[49] The Court will only consider overruling a policy when a state is alone or almost alone in certain ethical and moral choices.[50]

There is one final important way in which anti-discrimination law may preserve or even deepen social inequalities. This is through the very heart of the doctrine: states must discriminate on the basis of relevant factors. Therefore, in principle, if the police have credible information that a terrorist attack will be committed by a white female older than 65, it must focus its use of force on that group. There is a lot of discussion in the AI context about the self-reinforcing effect of group profiles and the loops that may arise. For example, a predictive policing system that suggests that the police monitor a certain neighbourhood with a significant immigrant population more than others will result in more data about crimes in that neighbourhood (and its inhabitants). This will lead to the AI system recommending stronger monitoring of that neighbourhood, and so forth. In a way, this is precisely what anti-discrimination law mandates. It requires good grounds for the police to monitor a certain neighbourhood, and a higher crime rate in that area may be exactly that. Courts generally do not require police units to assess the origins of social inequalities, let alone to correct them. They may sympathise with countries and organisations and use the data for constructive policies (e.g. investing more in education, social welfare or coaches to help people in those areas flourish personally and economically, which has a positive effect on crime rates) instead of attaching repressive consequences to these data. However, this is not a matter of non-discrimination law.

This is important because although many ethical frameworks for AI consider potential corrections for historical errors, this is not what non-discrimination law requires. Though in exceptional cases, positive discrimination is allowed or even deemed implicit in and thus required by Article 14 ECHR, there should be good grounds for doing so. The state should prove that there are factual inequalities. For example, when a country had favourable tax provisions in order to incite married women to work, this was not deemed to be in violation of the Convention:

> [T]he tax provisions which result in extra tax advantages accruing when a wife is the breadwinner of a family can be said to fall within the margin of appreciation

49 ECtHR, S.H. and others v. Austria, application no. 57813/00, 03 November 2011, § 97.

50 ECtHR, Fedotova and others v. Russia, application nos. 40792/10, 30538/14 and 43439/14, 17 January 2023.

> accorded to the national authorities. The Commission, therefore, finds that the difference in treatment in the present case has an objective and reasonable justification in the aim of providing positive discrimination in favour of married women who work.[51]

However, the Court critically assesses such policies. In principle, positive discrimination is treated the same as any form of discrimination. There should be good grounds for doing so, the effects should not be disproportionate to that aim and there should be no alternative non-discriminatory grounds available to achieve the same goal.

## 4. Approaches to Filling the Gaps

The analysed gaps are considerable and make it clear that the parameters for decision-making on discrimination are designed to require continual updating. This works in society, as long as discrimination can be identified and claims made. However, with regard to AI technologies and data-driven discrimination in general, the speed of evolution and the frequent opacity of the discriminatory logic make it hard for the same processes of claim-making to take place. As such, some important gaps are left by the current interpretation of anti-discrimination law amongst the member states of the ECHR. These gaps will be discussed in the present section.

### 4.1 Challenges

The justice system does not take structural discrimination (nor the perspective of groups) into account as a central issue in defining illegal discrimination and thus in enforcing anti-discrimination provisions. This leads to a system where redlining (providing different commercial services to people based on their location of residence) is permissible. It could be said that companies and by extension the court normalise and justify residential segregation by permitting this redlining. This is problematic because residential segregation has not only correlated but causally connected with inequal life chances based on structural factors, such as racism and poverty. To counter redlining, courts would have to start from the perspective that people have different access to opportunities based on historical discrimination and inborn or inherited characteristics.[52] As such, the use of proxies is not only not prohibited but highly likely, since almost

51 ECmHR, Lindsay v. UK, application no. 11089/84, 11 November 1986.

52 See, for example, the explanation provided by Binns, R. (2018, January). Fairness in machine learning: Lessons from political philosophy. In *Conference on fairness, accountability and transparency* (pp. 149-159). PMLR.

every social factor acts as a proxy for another. This makes it very hard to set out clear anti-discrimination rules or requirements for those creating models or systems, as it essentially requires the use of every variable to be justified negatively as non-discriminatory rather than providing a list of prohibited categories and chasing down each instance where one variable is used to proxy for another through legal claims.

A different problem is illustrated by predictive policing technologies. These technologies are based on two assumptions. First, that the definition of 'crime' used in these systems is appropriate to the goals of society in supporting law enforcement. This is not a given. For instance, in the US, the notion of 'drug crime' has been defined by a focus on crack cocaine and other drugs used primarily by the poor since the 1980s, producing an enforcement bias toward people of colour. One alternative would be to look at the trade and use of cocaine in general, which would then identify a much broader subsection of society as committing drug crime. Similarly, 'predictive policing' technologies tend to have an inbuilt definition of crime that focuses on offences committed by the poor (street violence, small-scale robbery and vandalism) rather than crimes per se. It would be equally possible to point the software at violence that occurs inside the home, such as child abuse or financial crime, which would provide a very different geographic rationale for both monitoring and enforcement.

These problems also apply to other systems, such as welfare benefits, child welfare interventions and student loan anti-fraud measures. For legal purposes, the problem is that neither the definition of crime nor the implicit bias in that definition are typically evaluated by a judge. Nor would a judge normally assess the extent to which a focus on certain crimes over others is legitimate. A judge would also not normally assess to what extent data gathering about crime is biased or how that bias influences the government's perception of the ground truth and the most appropriate action for it to take. These aspects normally fall under the margin of appreciation attributed to states. Moreover, an evaluation of whether governmental policies and actions are in the public interest is generally avoided by the European Court of Human Rights. This is already problematic, but these problems will be intensified by AI systems. One solution would be to have the judicial branch scrutinise these points in more detail, which would require it to adopt a bigger role in assessing the legitimacy of the dealings by the executive branch and to have an understanding of data management, quantitative research and statistics.

A second related assumption is that the police, the justice system and those who build technology for policing purposes have access to the ground truth about the prevalence of crime, and that historical records of arrests and convictions are statistically representative of all crime. This can be challenged on the basis that crimes committed in wealthier communities are less visible than those committed in poorer ones. For instance, they frequently occur indoors rather than outside, and richer people are exponentially less likely to be convicted of crimes (or to receive custodial sentences if con-

victed) than poorer people. Thus, our understanding of who commits crime is likely to be historically skewed toward the poor, and crime prediction algorithms, which work from this data, will tend to be pointed toward lower-income communities, which in many countries also have more minority residents. This leads to a situation where an algorithm devised to identify and predict 'crime' will actually identify and predict poverty or ethnic minority status, and where applying such an algorithm in policing will lead to the conflation of crime with these characteristics.

The legal paradigm is not capable of explicating what outputs of algorithms actually represent. To find a solution to this, the police (and more broadly, the government) should start from scratch and pay similar bias towards gathering, labelling and analysing data as is common in statistics and social sciences. This would mean either going through the tedious process of analysing what biases occur in the current dataset and correcting those in detail, or to start with the data collection process anew. This would mean that historical analysis on the basis of this less-biased dataset could only start in several years.

There is a further problem with the notion of 'factual inequalities' in relation to issues of ethnicity and racialisation. In most European countries, administrative statistics on categories other than 'western' and 'non-western' are not collected on the basis, ironically, of historical discrimination and violence.[53] This leads to situations where if someone has immigrant ancestors but is still subject to discrimination because of their skin colour or other attributes that persist over generations, it becomes difficult to prove that this discrimination is based on ethnic or racialised attributes because the person in question has become administratively 'western': they have been absorbed into the population and are no longer officially marked as 'other'.

Finally, there is a problem where discrimination tends to often be intersectional, or compounded across different attributes. Crenshaw's theory of intersectional discrimination[54] holds that injustice occurs where someone has to pick an attribute on which to make a claim of discrimination, when in fact discrimination occurs based on intersecting characteristics, for example being a woman of colour, or being disabled and transgender. The idea of the comparator becomes problematic here in a different way, since the group of those that share these intersecting characteristics will necessarily be much smaller than a grouping based on just one of them. Here, court convenience supersedes individual needs.

53 Bonjour, S. (2020). Epilogue: Shaping the Nation through civic integration: A postcolonial perspective on paradoxical policies. *Revue europeenne des migrations internationales*, *36*(4), 135-142.

54 Crenshaw, K. W. (2013). Mapping the margins: Intersectionality, identity politics, and violence against women of color. In *The public nature of private violence* (pp. 93-118). Routledge.

### 4.2 Responses to these Problems

One possible response to the gaps described above is, of course, for human rights law to evolve so that it can take account of them. However, such change occurs very slowly over time. An illustration of this is the 1979 decision, cited above, which defined the nuclear family as a moral good. Courts have moved slowly to recognise other types of family beyond the nuclear, married male and female plus children configuration, for instance to recognise gay couples or transgender people as being deserving of protection. Yet they have not moved forward in recognising the kind of problem Crenshaw identified with the compounding nature of discrimination based on gender plus racialised or ethnic characteristics, or other combinations of attributes that have become relevant through public debates over the decades. As such, there is little evidence that courts are willing to provide a systematic or potentially intersectional way of dealing with discrimination – or even to countenance the conditions under which one might bring such claims in the first place.

A different response in the field of algorithmic discrimination aims to close the gaps from the technological side. This response has come from the dialogue between computer scientists and social scientists on fairness in relation to machine learning and automation. Here, the American idea of fairness as the prevention of disparate impact[55] has been highly influential due to the establishment of academic fora for this discussion, initially in the US. Despite the move away from US framings toward increasingly philosophically-informed debates on the possibilities for understanding discrimination and unfairness in different ways,[56] this thinking has largely failed to move European legal debates toward considerations of intersectionality or the rights of groups who may find themselves discriminated against as groups on the basis of algorithmically-selected characteristics that do not follow the lines of traditionally-defined discriminatory categorisations. One reason why there is less receptivity to this approach may relate to the prohibition on gathering and storing data on ethnic and racialised backgrounds or characteristics. In contrast, in the US the wealth of statistical data collected by the state, employers and institutions on these demographic characteristics makes it easier to demonstrate quantitatively disparate impact, providing a more compatible starting point for anti-discrimination thinking in relation to technology.

Another solution might come from self-regulation and standardisation. For a period of time, many ethical choices that conflicted with the commercial interests of the Big Five were demanded and implemented through pressure from their workforce. Now,

55 Barocas, S., & Selbst, A. D. (2016). Big data's disparate impact. *Calif. L. Rev.*, *104*, 671.

56 Binns, R. (2018, January). Fairness in machine learning: Lessons from political philosophy. In *Conference on fairness, accountability and transparency* (pp. 149-159). PMLR.

with the cramped job market, their power is waning, and enterprises are reducing their focus on and investment in ethics, due diligence and duties of care. At the same time, established companies are eager to maintain their monopolies and actively push for legislation that places administrative burdens on AI developers. To the extent that discrimination is equivalent to the use of irrelevant criteria, there is an interesting alignment of the interests of the users of predictive algorithms and the citizens that are affected by them. Both want a system that is geared towards using the most relevant and adequate factors and one that is based on a correct representation of societal reality. Only the developers of algorithmic systems have an incentive to not do so, because building such systems is more difficult and painstaking than building based on pseudo-science, which is true for almost all of the existing systems.

This stream of computational thinking about discrimination has given rise to a set of technical ex-ante approaches to discrimination through AI technologies. One is the widespread preoccupation with 'debiasing AI',[57] which denotes the scrutiny and pruning of training datasets or the behaviour of models to make them less likely to inform discriminatory decisions. However, this approach has been criticised on two fronts. First, it promotes a temporary and limited fix for problems that are systematic. After all, data will continue to be created under conditions of structural inequity, making it impossible to erase discrimination from the data. Second, infrastructures and models are themselves biased, and cleaning specific datasets will not remedy that problem in a meaningful way.[58] This critique could also be levelled at many of the technology-centred guidelines that have been produced to help computing sciences counter bias in models and data, namely the FACT principles (Fair, Accurate, Confidential and Transparent), and approaches using synthetic datasets. Since the world is biased, scholars of discrimination and AI ethics argue, it is hard to imagine how to create either datasets or models that can be used in the world but do not reflect and reproduce that bias. One prominent account of this problem is given by Bender et al.[59] in their paper on large language models being developed by Google: these models are trained on gigantic quantities of natural language from the internet to 'learn' how to produce natural language responses to prompts. However, the language available on the internet reflects a world of inequity

57 Majumdar, P., Singh, R., & Vatsa, M. (2021). Attention aware debiasing for unbiased model prediction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 4133-4141). Deshpande, K. V., Pan, S., & Foulds, J. R. (2020, July). Mitigating demographic Bias in AI-based resume filtering. In *Adjunct publication of the 28th ACM conference on user modeling, adaptation and personalization* (pp. 268-275).'

58 <https://edri.org/wp-content/uploads/2021/09/EDRi_Beyond-Debiasing-Report_Online.pdf>

59 Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

and bias, meaning it is impossible to build such a model that will not replicate the use of language and concepts common in online spaces.

Some approaches have argued for a broader perspective to be adopted on mitigating discrimination, by not only focusing on the bias in existing data sets and models but on the broader sociotechnical system in which technologies are developed and used.[60] As we have seen, bias can be introduced in how technology is used or how problems are framed. Such approaches emphasise the need to also address biases in organisational and governance structures, by for example including affected groups in the problem formulation and the design of systems.

The lack of response to problems of intersectionality in discussions of AI is an issue of both depth and breadth. Two factors are relevant here. First, the ability of those building models and curating training data to label multiple characteristics as sensitive, in combination with each other as well as on their own. This is not technically impossible, but it is an engineering issue that would require common language and procedures currently underdeveloped in the field of AI. The second factor relates to agreeing what factors should be considered sensitive in combination, and how those combinations should be treated computationally. Sensitive or prohibited categories are currently identified through law, that is, those categories have been legally prohibited or decided to be impermissible through jurisprudence. Intersectionality still presents the same bottleneck for legal decision-making that Crenshaw described in 1989. Until a combination of attributes is formally labelled as sensitive or unusable through legal decision-making, it will not become a decision-making constraint for computer or data science. Although intersectional problems are increasingly a common topic for ethics discussions, they have still not found a functional route into policy or legal decision-making, three decades after they were formally identified as a legal problem. For them to become embedded in computer or data scientific practice, these problems seem to need a higher profile. Ironically, the nature of intersectional problems – as it pertains to smaller groups than traditionally recognised problems of discrimination – also leads to their being under-prioritised, making it less likely they will be addressed as an issue in computer science in the near future.

Although the discussed approaches do not address the gaps in discrimination law directly, they can nevertheless inform regulatory initiatives, for instance through standardisation. Standardisation and regulation are part of the response to discrimination through AI technologies. Human rights impact assessments, for example, could be part of a standardisation or regulation approach, depending on how they are used. These have been discussed as a way of broadening the set of rights under consideration when

60 van der Sloot, B., Keymolen, E., Noorman, M., Weerts, H. J. P., Wagensveld, Y., & Visser, B. (2021). Handreiking non-discriminatie by design.

it comes to AI. These have been proposed in the US[61] and Canada,[62] and they are being recommended in the UK and the Netherlands on the same basis.[63] They can help to broaden the scope of mechanisms to address discrimination in the development and use of technologies, such as by also looking at the governance structures around the technology. However, these impact assessments require a readiness to define problems as discrimination when they affect groups, and they may result in challenging developers and courts with other forms of discrimination than those already catalogued in legal decision-making.

Within the regulations, fundamental rights are cited in ex-ante approaches, such as the EU's draft of the AI Act.[64] This might prove problematic in terms of identifying actual instances of discrimination for several reasons. First, like ECHR human rights law, a fundamental rights focus distances groups from the ability to make claims where AI is discriminatory, instead focusing on individuals as the recipients of discriminatory treatment. However, the proposed regulation does not provide mechanisms for either groups or individuals to make claims. Consequently, the fundamental rights focus becomes a general statement about building rights-compliant AI rather than enabling affected people to respond to discrimination or other rights violations. The Act does name vulnerable groups whose interests should be especially guarded with respect to AI applications (children, disabled people, those dependent on welfare, and migrants, among others) and thus provides language for thinking about groups.

However, the draft Act then falls back on existing anti-discrimination law for the definition of what might be problematic with respect to AI's effects on those groups. For example, with respect to the public sector and social welfare, the Act states:

> If AI systems are used for determining whether such benefits and services should be denied, reduced, revoked or reclaimed by authorities, they may have a significant impact on persons' livelihood and may infringe their fundamental rights, such as the right to social protection, non-discrimination, human dignity or an effective remedy. Those systems should therefore be classified as high-risk.[65]

61 Moss, E., Watkins, E. A., Singh, R., Elish, M. C., & Metcalf, J. (2021). Assembling accountability: algorithmic impact assessment for the public interest. *Available at SSRN 3877437*.

62 Yam, J., & Skorburg, J. A. (2021). From human resources to human rights: Impact assessments for hiring algorithms. *Ethics and Information Technology*, *23*(4), 611-623.

63 Edwards, L. (2022). Regulating AI in Europe: four problems and four solutions. *Ada Lovelace Institute*, *15*, 2022.

64 Proposal for a regulation of the European parliament and of the council laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts, com/2021/206 final.

65 Recital 37 Proposed AI Act.

However, this defines discrimination as something that can be identified by the designer or deployer of a system before that system is in use, which is a necessary but insufficient component of a strategy to define and prevent potential discrimination.

Related to this, counterfactual approaches to identifying discrimination have been advocated as a way to provide clarity and certainty on what is going wrong in models that may otherwise be opaque. Wachter et al. argue that testing a model using counterfactual inputs to force it to give alternative decisions, then studying how those differ from those made in relation to a particular data subject, 'provides data subjects with meaningful explanations to understand a given decision, grounds to contest it, and advice on how the data subject can change his or her behaviour or situation to possibly receive a desired decision (e.g. loan approval) in the future'. However, the authors note that these constitute 'a minimal form of explanation',[66] and that:

> where it is important to understand system functionality, or the rationale of an automated decision, counterfactuals may be insufficient in themselves. Further, counterfactuals do not provide the statistical evidence needed to assess algorithms for fairness or racial bias.

As such, counterfactuals cannot help with cases where statistical proof is needed, and they also necessarily overlook the structural aspects of discrimination because they focus on individual cases. Further, they are unlikely to provide useful conclusions where discrimination cannot be reduced to a single attribute (i.e. in cases of intersectional discrimination).

With respect to AI technologies, this set of possibilities leaves us in a bind. If law does not define the use of certain categories as discriminatory, it seems we may need to start from bans and prohibitions or possibilities for making claims in non-tech-related environments.

## 5. What Are Possible Ways Forward?

Starting with a freedom as non-domination perspective on discrimination, this chapter has analysed several of the fault lines AI has made visible in how anti-discrimination law and regulations address discrimination. The first point is they cannot take account of intersectional forms of discrimination. Intersectionality is and will continue to be a major challenge to both legal and computational approaches to fairness. Until it is meaningfully tackled, it cannot be said that fairness has been addressed adequately.

66 Wachter, S., Mittelstadt, B., & Russell, C. (2017). Counterfactual explanations without opening the black box: Automated decisions and the GDPR. *Harv. JL & Tech.*, *31*, 841, p. 883.

However, tackling it also entails naming intersectional discrimination as a problem, analysing its dimensions and implications, and recognising those who are being affected by such forms of discrimination. All of these are social and political questions, and in turn drive legal questions. The computational understanding of intersectionality as a problem of discrimination will only open up once the social dimensions are more fully acknowledged and named. Once it has become clearer which intersectional problems of discrimination should be addressed, it is possible to engage with them on the computational level. It is feasible to create datasets and design AI models in ways that make it possible to reduce intersectional discrimination. However, first, from a regulatory perspective, it is necessary to provide an incentive to do so by requiring that those building statistical models pay attention to multiple sensitive attributes (or proxies for them) simultaneously. At the same time, from a practical perspective, this must be done by mandating the creation of cross-tabulated variables composed of attributes that, together, form sensitive or protected categories (e.g. gender and ethnicity) in order to make it possible to do so. This potentially reduces the analytical power of the resulting AI models but is feasible in design terms. This approach has not been tested because discrimination law has not demanded it until now.

The second problem we identify from the freedom as non-domination perspective is the question of what constitutes arbitrariness. Arbitrariness, from a legal perspective, arises when there are no rational criteria for the use of power or objective grounds to use that power. While focusing on certain ethnicities when patrolling borders is commonly regarded as arbitrary from a societal perspective, it may not necessarily be so from a legal perspective, and it is therefore unlikely to be considered arbitrary by a court. AI adds an extra layer to this uncertainty, since with sufficiently large datasets a correlative grounding of any use of power can almost always be found. Arbitrariness in the exercise of power through AI models is therefore not defined by whether there is a correlation between a proxy and a determinative, but by arbitrariness in how data are collected and analysed, and in how correlations are interpreted. These types of questions are difficult to address under the current legal paradigm because of the lack of nuance in the law's definition of arbitrariness.

Moreover, it is particularly on this point, and the related point of effectiveness – does this AI system help in fighting crime, for example, or does using this proxy help in combatting terrorism – that judges allow states a particularly wide margin of discretion. Suppose one could reasonably predict who has a high potential for becoming criminal, and one could come up with the 400 most likely criminal masterminds in a city. What is the best way to use that information: by investing in pre-emptive policing or by investing in education, social services, and welfare assistance offered to families and neighbourhoods considered prone to produce criminals? These questions are seldom discussed by judges, most likely because this is seen as the job of social policy only.

Moreover, the self-reinforcing logic characteristic of pre-emptive policing (that those who are policed most are most likely to be caught committing infractions)[67] is seldom taken into account.

This leads to the third and final point: the divide between legal scholars and judges' understanding of how decisions should be justified, and that of data science and AI experts. For the former, decisions must be justified with reference to the body of law and jurisprudence. Comparatively, for the latter, justification lies in the statistical confirmation of a hypothesis. This constitutes a fault-line because it is possible to use data-driven analysis to demonstrate hypotheses that have no grounding in science or reality, and which are patently false. The AI Act and AI Liability Act may help, inter alia because the latter spells out that organisations must be open about their data governance when liability issues arise. However, human rights law violations are substantially different than questions of liability. The uncertainties that result from the use of data-driven AI technologies, in particular opaque ones, increase the risk of harm to particular groups in the form of exclusion and infringements on their freedom, for example on their freedom of movement. Such uncertainties that carry through in judicial decision-making weaken the checks and balances and reduce the extent to which affected groups can have a say in or object to the rules that they are subjected to.

One important if inadvertent contribution to the debate on AI and discrimination has been made by the realisation that discussions on how to formalise 'fairness' into computational models have not made the progress that was hoped for a decade ago. Instead, these discussions have become more intense and interdisciplinary as AI systems have become more commonplace in society, without achieving the requirements that were originally the objective of this process. What has been achieved, however, is a socio-scientific and societal debate that has deepened the understanding of just how large the shortfall is between what can be formalised into models and systems, and the complexity of the notion of fairness in relation to our interactions with AI and data technologies. Computer science debates are still missing the inclusion of many important aspects of fairness, especially those recognising historical disadvantages and current disempowerment.[68] As such, fairness is inherently a moving target – a characteristic that clashes with most legal and computational approaches that attempt to pin it down.

Promising approaches that have come out of this debate include the drive to define discrimination as more than bias; something that can be isolated and identified as a conscious choice on how to build a model or a dataset. Moreover, the literature shows

67 See, for example, an alternative mapping of crime in New York City focusing on white-collar infractions: <https://whitecollar.thenewinquiry.com/>.

68 Selbst, A. D., Boyd, D., Friedler, S. A., Venkatasubramanian, S., & Vertesi, J. (2019, January). Fairness and abstraction in sociotechnical systems. In *Proceedings of the conference on fairness, accountability, and transparency* (pp. 59-68).

a gradual movement toward the attempt to understand how factors become discriminatory in their use, and toward what can and cannot be applied to the public, and how. One branch of this type of thinking is visible in civil society organisations, which are often charged with defining how groups should make claims about discriminatory systems.[69] However, it also appears that AI systems can be discriminatory in ways that are hard to predict in academic literature on discriminatory and harmful AI and in work on optimisation.[70]

One more promising direction is to not see algorithmic discrimination as a binary problem of either a technology or a societal issue, but as sociotechnical problem that requires a broader view on how individuals or groups can be inhibited in their ability to participate equally. This requires ex-post and ex-ante measures because AI constantly creates new kinds of uncertainty about how it will be used.

The unsolved problems of AI and discrimination also point to the issue of power at scale. Infrastructural power[71] is a broader issue than discrimination, but it helps to explain why discrimination is not being tackled effectively. It is institutions and companies that build and manage the infrastructures through which data is transformed into the power to intervene, and the power these parties exert is of a scale that, unless regulated, tends to produce political domination.[72] In turn, this domination is the overarching force that restricts the freedom of groups, and discrimination tends to play a role in achieving this. Much has been written about this problem in the field of privacy[73] and discrimination studies. However, these two sets of arguments have not been brought together effectively as a critique of infrastructural and algorithmic power as of yet. Where AI technologies combine with data-driven analysis to provide tools for domination, they can do so by exposing people's behaviour and attributes to make them tools for either political or market intervention – through a general form of domination, to put it differently. They can also do so by providing ways to exclude or exploit based on particular histories or group attributes,[74] by employing discrimination as a tool of dom-

69 <https://edri.org/wp-content/uploads/2021/09/EDRi_Beyond-Debiasing-Report_Online.pdf>.

70 Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

71 Gurses, S., & Van Hoboken, J. (2017). Privacy after the agile turn.

72 van der Sloot, B. (2018). A new approach to the right to privacy, or how the European Court of Human Rights embraced the non-domination principle. *Computer law & security review*, *34*(3), 539-549.

73 Roberts, A. (2018). Why Privacy and Domination? *Eur. Data Prot. L. Rev.*, *4*, 5. Newell, B. C. (2014). Technopolicing, surveillance, and citizen oversight: A neorepublican theory of liberty and information control. *Government Information Quarterly*, *31*(3), 421-431.

74 Gebru, T. (2019). Oxford handbook on AI ethics book chapter on race and gender. *arXiv preprint arXiv:1908.06165*. Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. John Wiley & Sons. Eubanks, V. (2018). *Automating inequality: How high-tech tools profile, police, and punish the poor*. St. Martin's Press.

ination. The former addresses people as personal data to be harvested,[75] while the latter produces power over populations by dividing and categorising, and most significantly by leveraging the weight of historical injustice.

This analysis could provide a way to reframe the problem of AI and discrimination as requiring new legal doctrine. Instead of using legal reasoning devised for analogue problems to mitigate technological and infrastructural domination, what would it mean to begin from the problems of technology? This would require new legal reasoning that addresses problems of domination as they are created by both corporate and public technological architectures; problems of making claims in relation to our identities as groups, as individuals experiencing intersecting and interacting forms of discrimination, or as subjects of pervasive technological surveillance and sorting. It would require rethinking the enforcement of existing laws and regulations, but also reframing those under development. Such an effort would also have the effect of reframing the ways in which we can address digital technologies, and AI specifically, as forces that define societal and economic aims, and the extent to which we are willing to accept a determinist view of such technology as influencing what is possible and what is legitimate, instead of first articulating these boundaries and shaping technology within them.

75 Cohen, J. E. (2018). The biopolitical public domain: The legal construction of the surveillance economy. *Philosophy & Technology*, *31*, 213-233.

CHAPTER

# XII

# What Is Law? A Fuller Perspective on Legal AI

**Bart van der Sloot**[1]

https://doi.org/10.26116/ewbr-8w91

1 Associate Professor, Tilburg Institute of Law, Technology, and the Society (TILT), Tilburg Law School (TLS), Tilburg University.

## 1. Introduction

There have been many experiments with automating the legal regime at its most basic level and using artificial intelligence (AI) to assist human decision-making. This can be done through, inter alia, providing lawyers, judges and lawmakers with relevant case law, evaluating the validity of the arguments presented by the various parties, writing the non-operative part of a judgment (i.e. providing a summary of the arguments of the parties), reciting relevant laws and case law, giving a description of the relevant facts and even (although this lies at the border of the second approach) providing judges with draft verdicts, which they would need to verify, amend and certify before it could be finalised. In addition, AI is deployed to replace human decision-making. In this scenario, AI is instrumentalised to do what humans would do in the manner they would go about it.

More audacious still is the effort to use AI to perfect the legal regime, for example by trying to make it more rational and objective. Under the current legal regime, laws are not always internally consistent, terms and definitions are used in varied ways and some laws and policies have opposite effects. The quality of judgments not only suffers from flaws, judges' personal preferences and their individual convictions, but also from the conditions under which these judges delivered the judgment (were they hungry, were they under time-pressure, were they preoccupied with a private matter?). A promise of AI is that it can take away those contingencies, and issue judgments on objective and consistent criteria, so that laws can be made more rational, objective and consistent in terms of language, range and effect.

There are many concerns about and criticisms of legal automation, such as whether AI will ever deliver on its promise, or with respect to AI's inherent bias and unforeseen side effects. Suppose, however, that AI could deliver on its most far-reaching promise and could not only assist human decision-making or replace it but perfect it. What would the legal regime look like and is that what a legal regime should look like? The answer to this question, of course, depends heavily on the philosophical understanding of what the legal regime essentially is and should be. It is remarkable that although there is much discussion about the effectiveness of AI, how AI systems work in practice and what technical qualities AI needs to improve on, there are relatively few reflections on what a 'perfect' legal regime looks like. To the extent there is a silent but underlying understanding of the law that drives the current approach to legal AI, it is an understanding of law that aligns with what legal philosophers would call legal positivism.

In essence, the dream of legal positivism is to arrive at a legal system that is consistent, rational and objective. Law is that which is posited: nothing less, nothing more. It is a hierarchical system, in which the *grundnorm* or the set secondary rules have the authority to enact laws and to decide how this should be done. Taking a factual approach

stripped of moral sentiments, legal positivism is resistant to customary law, intuitions and supra-legal moral norms. With leading scholars such as Austin, Bentham, Hart and Kelsen, legal positivism conceives the legal order as a closed organism that is self-referential; laws should be formulated as unambiguously and as consistently as possible. Within legal positivism, there is a focus on the legislative branch; although the judicial and executive power have some margin of appreciation, in the ideal version of law, the prerogative lies with the legislative power, adopting laws that are so clear, precise and detailed that the judicial and executive power can simply apply the laws in practice and to concrete situations.[2]

In theory, legal AI could perfect this approach to law. Computer programs allow for a consistent and hierarchically organised set of rules. Computer language is also less ambiguous than natural language and is devoid of subtext or supra-textual intention. It can codify clear decision trees, with a potentially endless set of criteria and sub-criteria programmed into it. AI (the non-deep learning type) can also make legal reasoning and argumentation explicit, visible and thus verifiable and auditable, given that the more opaque and mystic way in which judges tend to arrive at conclusions forms a daunting problem for legal positivists. In AI systems, codified rules lead directly to outcomes when fed with casuistic data, so there is no need for the discretion currently attributed to the executive and judicial power.

A second commonly referred to philosophy of law is natural law theory. This understanding of law is similar to legal positivism in that it is hierarchical. In its classical form, the highest norm is not derived from human authority, the *grundnorm* or secondary rules, but from God. God is seen as the ultimate lawgiver, and the representative of God (e.g. the King) is deemed to hold the supreme authority to issue laws. In its more modern variant, natural law theory underlines the existence of pre-legal and supra-legal norms. Natural laws, like gravity and thermodynamics, are eternal, stable and cannot be violated: they codify the basic conditions for human life. Where legal positivists separate is (is this a law?) from ought (is this a just law?), under natural law theory, these questions are marginally linked (if the law is so bad that it tramples basic natural rights, it is not a law or the law should not be effectuated). What those pre-legal and supra-legal norms entail depends on which natural law theory is consulted, but a commonly-shared minimum principle is that a law that would undermine citizens' human dignity or lead to the annihilation of certain groups in society would be null and void. Natural law theory views the legal order as aspirational or teleological; law aims at the best, morally good community to come, although no law will ever be able to deliver on this ideal in full.

2 To the extent that such is not possible, discretion for the judiciary and executive power is seen as a necessary evil rather than a virtue of the legal regime.

Although there are some aspects of natural law theory that would merit legal automation, the match is less perfect than with legal positivism. To the extent that pre-juridical norms and moral limits can be explicated, AI systems could codify them and potentially perform a final assessment before making a decision, checking whether the decision would violate the minimum moral standards. However, it is clear that at least three obstacles emerge. First, there is no consensus or clarity about which moral limits are to be considered natural rights. Second, minimum norms such as 'human dignity' are so vague and abstract that it is unclear how an AI-system could use them as benchmarks. Third, oftentimes an evaluation would depend on the actual consequences of laws, policies and judgments, even though they may not be clear or foreseeable when the AI system would need to make a determination. On other points, however, natural law theory differs little from legal positivism. Like legal positivism, for example, it is based on a highly hierarchical system and aims to achieve consistency and unambiguity. Thus, although not a perfect match, legal AI has both merits and pitfalls when law is considered from a natural law perspective.

It is only when a third understanding of law, arguably the most appealing, is adopted that legal automation becomes truly problematic. This approach is known variably as sociology of law, legal pragmatism and legal empiricism. What they have in common is that, essentially different from both legal positivism and natural law theory, they treat law as a non-hierarchical, bottom-up human construct. This understanding of law ultimately denies that law is or should be consistent and unambiguous. 'Is' and 'ought' cannot be separated here; rather, to understand what a law is requires an understanding of what the laws aims at. In many cases, there is neither a legally nor a morally satisfying answer; this is why non-legal forms of dispute resolution and non-legal forms of regulation are valued highly. Under this understanding of law, it is difficult to see how law can be automated or put differently. If law were to be automated, this understanding of law would find that the essence of a legal regime proper would be corrupted.

This chapter explains this point in more detail by showing how fundamentally opposed legal AI is to the third understanding of law. It does so by homing in on the work of one of its most reputable proponents: Lon Fuller. The main body of this chapter lays out the most important aspects of his philosophy, which may be interesting to lawyers and regulators in their own right because his theory of law is arguably one of the most superior, although not well-known. Section 2 discusses Fuller's understanding of the origins of the legal order. Section 3 assesses what is required for a legal system to be effective and legitimate and for it to attain its raison d'être. Section 4 evaluates the essential characteristics of a legal order under Fuller's approach to law. Because most efforts in legal AI are currently focussed on assisting, replacing and perfecting jurisprudence, section 5 provides Fuller's take on case law and the proper role of the judiciary. Section 6, finally, considers what Fuller's work implies for the effectiveness and desirability of legal AI.

## 2. The Origins of Law

This section discusses the foundations of the legal order as conceived by the third philosophy of law by giving three illustrations from Lon Fuller's work. First, in contrast to legal positivists who see law as a closed and self-referential system, Fuller underlines that the author of the law, their intentions and the context in which the law was adopted are not only visible; they are essential to understand the law (Section 2.1). Second, in contrast to the other two philosophies, law is not conceived as something posited by a human or godly lawgiver, so much as they are reflections and codifications of customs and social norms (Section 2.2). Furthermore, rather than taking a factual and textual approach to law, Fuller underlines the essence of law as a symbolic order (Section 2.3). Finally, a small wrap-up is provided (Section 2.4).

### 2.1 Implicit Laws

In his short story 'My debut as a literary person', Mark Twain recounts the tale of a boat crew stranded on a desolate spot of land. After a few days, the men become desperate as proper food is lacking, and the captain and the two passengers start scraping boot-leather and wood, then make a pulp of the scrapings and moisten them with water. The sailors, however, do not make pulp but start to eat strips of leather from old boots, with chips from the butter cask. When one of the crew mates was asked about the affair afterwards, he remembered that the boots were old and full of holes and, he added thoughtfully, that it was the holes that digested the best.[3] It is to stories like this that legal positivists jokingly refer when describing the position of natural law theorists: they always prefer the gaps over the legal order itself.[4] That is, natural law theorists attach more weight to non-codified pre-legal norms than to codified legal norms. Legal positivists, by contrast, believe that the optimum is a closed legal system.

Fuller disagrees with both positions. He believes that there are aspects that law cannot account for but that these are not higher pre-legal or supra-legal standards. Law is a human product and, as such, has intrinsic limitations.[5] It is not that the holes precede and supersede the legal order, but rather that there are practical restraints on what a

3 Twain, M. 'My debut as a literary person' <https://www.public-domain-content.com/books/Mark_Twains_Short_Stories_2/C6P7.shtml>.

4 See for example: Somló, F. (1917). 'Juristische Grundlehre', Leipzig, Felix Meiner, p. 410.

5 Customary law, the basis of the whole legal regime, is also implicit to Fuller: 'Customary law can be viewed as being implicit law in a double sense. In the first place, the rules of customary law are not first brought into being and then projected upon the conduct they are intended to regulate. They find their implicit expression in the conduct itself. In the second place, the *purpose* of such rules never comes to explicit expression.' Fuller, L. L. (1971), 'Anatomy of the law', Penguin Books, Harmondsworth, p. 44.

lawmaker can achieve, similar to practical limits for a locksmith, a lumberjack or a carpenter. A fully closed, self-referential system, such as legal positivists aspire to, is simply consequently impossible:

- First, law needs to be general, because it would be impossible to codify every specific situation that might arise. 'If the draftsman were to attempt to forestall in advance every conceivable aberration of the legislative power, his constitution would resemble a museum of freaks and monsters. It is certainly difficult to imagine such a constitution serving an educational function or offering a suitable object for a pledge of allegiance.'[6]
- Second, because it is impossible to foresee all future scenarios, not only does the law need to be formulated generally, but it should also be interpreted according to the legislator's intentions. A lawmaker, for example, could not have made explicit that a consulship could not be held by a horse because this was simply unthought of.
- Third, not only can law not account for every potential future scenario, but by necessity it has to be tailored to the normal course of affairs, rather than the exceptions (*necessitas non habet legem*).
- Fourth, law cannot account for situations in which it is dispensed with, such as after a regime change, coup or revolution. 'What laws now govern events that occurred during the six months while the rebels held the reins of government? Do we make bastards of all the children of marriages contracted under new marriage laws enacted by the intervening regime and now repealed by the returning government? If we declare these marriages legally valid, do we also uphold every confiscation of property accomplished under laws enacted by the now deposed regime?'[7]
- Fifth, a legal regime cannot account for its own coming into existence: it cannot legitimise its own legitimacy.

## 2.2 Customs

Following from this practical account of the limitations on the lawmaker, Fuller emphasises that the legal order is a codification of pre-existing societal norms and customs. Accordingly, there are not only intra-legal limitations; there are also extra-legal foundations of the legal order.

6 Fuller, L. L. The Implicit Laws of Lawmaking, p. 178. In: Fuller, L. L. (1981). 'The Principles of Social Order', Duke University Press, Durham.

7 Ibid, p. 184

Contracts, for example, serve a transactional purpose between two private parties. For example, A agrees to let B use her land, and B agrees to pay A one gold coin each month. No government needs to exist for such a contract to be concluded, yet an ultimate arbiter should have the authority to decide over disputes, should they arise. Although parties can contractually agree on an arbiter, in most pre-modern societies, the nature of this arbiter was clear and did not merit codification (e.g. the head of the tribe, the priest or any other authoritative figure, depending on the community). From the social practice of drawing up contracts and the body of jurisprudence that emerged from disputes and settlements, a set of rules and principles on drawing up contracts materialised regarding, for example, which conditions should be included, which terms and agreements are null and void, and under which conditions and between what types of parties contracts can be concluded (consider rules on minors and legally incompetent persons, rules on deceit and abuse of circumstances). These rules and principles were later codified in a formal law adopted by the state. Consequently, contract law is, according to Fuller, historically the first branch of statutory law.[8]

The fact that in pre-modern societies often no arbiter needed to be appointed, means that customs are also important for attributing authority. Authority to set rules and issue judgments is an extra-legal element of a legal order, as legal pragmatists would point out, since law cannot give the lawgiver the authority the give law. Interestingly, Fuller points out, even legal positivists cannot avoid this point. Austin, for example, accepted that:

> the sovereign is that person or group of persons which society is in habit of obeying. It is clear that this conception rests the legal order ultimately on custom, since the sovereign is, under this view, merely the beneficiary of a custom of obedience, and the security of his position will depend upon the strength of the custom supporting it. It is equally obvious that in reality the role of custom is not limited to determining the sovereign power, but that the content of the law itself, and the allocation of powers among the sovereign's agents, are also in large part determined by custom. It is furthermore clear that it may be custom rather than the sovereign power which furnishes the basic stability of a society.[9]

8 Fuller, L. L., Eisenberg, M. A., & Gergen, M. P. (1981). *Basic contract law*, West.

9 Fuller, L.L. (2012). Law in Quest of Itself, The Lawbook Exchange, Clark, New Jersey, p. 31-32. See a similar point as to why citizens should follow the law: Fuller, L.L. (1953). American Legal Philosophy at Mid-Century--A Review of Edwin W. Patterson's Jurisprudence, Men and Ideas of the Law, 6 J. Legal Educ. 457, p. 468.

Customary law, also called 'the language of interaction' by Fuller, can be seen as the 'inarticulate older brother of contract'.[10] Contracts generally formalised and standardised a pre-existing norm or custom. A let B use her land and B paid A in return. Because contractual terms are fixed, a contract replaces trust with certainty and flexibility with consistency. Although it is often suggested that it is only when society becomes more complex that more parties get involved in legal relationships and when parties are more indirectly and distantly related to each other that it becomes necessary to formalise customs in contracts, the opposite is true. In pre-modern societies, a contract was not so much unnecessary as it was impossible. The pre-modern society is essentially a web of interconnected interdependencies.

> The members of a small, self-sufficient group are all parts, one of another; all are bound together by a complex network of reciprocal renditions and expectations. In such a human situation any attempt at an explicit verbalized definition of each party's expected performance, and the price to be paid him for it, would certainly not produce order and might produce chaos.[11]

Contracts should be seen as a move away from complex interdependencies (the mother of A once saved the brother of B when he was in jeopardy, but the son of B gave a gift to A's wife last month, and that is why A…) and towards a simpler, more atomistic society.[12]

Contracts are based on the customs and social understandings they formalise, which is why they need to be taken into account when interpreting the text of a contract. As such, contract law:

> stands halfway between customary law and enacted law, sharing some of the qualities of both. On the one hand, contractual law is like customary law in that its prescriptions are not imposed on the parties by some outside authority; they make their own law. On the other hand, contractual law resembles legislation in that it involves the explicit creation of verbalized rules for the governance of the parties' relationship.[13]

The fact that written/statutory law originated as a law of contracts and that contracts are a codified form of customary law also means that the legal regime is preceded by and based on individual autonomy (the desire and capacity of individuals to pursue

10 Fuller, L.L., The Role of Contract, p. 194. In: Fuller, L. L. (1981). 'The Principles of Social Order', Duke University Press, Durham.

11 Ibid, p. 200.

12 Fuller, L.L., Forms and Limits of Adjudication, p. 120. In: Fuller, L. L. (1981) 'The Principles of Social Order', Duke University Press, Durham.

13 Fuller, L. L. (1969). Human interaction and the law. *Am. J. Juris.*, *14*, 1.

certain goals and to attain those goals) and social reciprocity (the desire and capacity of two or more individuals to engage with each other in inter-connected and inter-depended ways).[14]

### 2.3 Symbolism

Next to intra-legal limitations and pre-legal foundations of the legal order, there are supra-legal elements created through the legal order. Legal positivism aims at a legal order that is devoid of ambiguity and vagueness, which is one of the reasons it is so opposed to the use of legal fictions. Bentham, for example, stresses that in 'English law, fiction is a syphilis, which runs in every vein, and carries into every part of the system the principle of rottenness'. Fuller points out that legal orders are reliant on legal fictions to a large extent: after marriage, a couple is treated as a single economic unit; liability law often attributes responsibility, for example holding that all traffic accidents between a motor vehicle and a pedestrian will deemed to be caused by the driver of the motor vehicle; an adopted child is legally treated as if it were the biological child of the adopting couple; if a person has gone missing and has not surfaced after twenty years, they are presumed dead; judges often adopt an 'it must be presumed that' type of reasoning; law creates 'legal persons'; there is the legal presumption that citizens know the law, and so forth.

Obviously, a company is not a person, there will never be a society in which all people know every law, and spouses are not merged into one being. Legal fictions, for Fuller, are not an inconvenience that need to be smoothed out: they are an essential part of the legal order. The legal order is not only a to be factual order, it is also a normative order, a description of what should be but may never be, borne from an aspiration that will never be achieved in full but that still should be the aspiration.[15] What is more, the legal order is formulated through language and, as such, depends on the use of symbols, concepts and abstraction. Thus, law is based on a symbolic language, which depends on abstractions and creates metaphors and symbols that capture how we want the world to be, even though it is not factually so.

14 Judges will still take standard practices in the sector in which the contract was concluded into account when interpreting contractual terms. Even though A and B might not have agreed explicitly that when A would deliver 10 barrels of milk each week, she would take with her the empty barrels from the previous week, a judge may hold A to that task if that is a standard practice in the milk delivery sector. B presumed and was entitled to presume that A would do so, without making such explicit.

15 Fuller, L.L. (1967). Legal Fictions, Stanford University Press.

### 2.4 Conclusion

Should this approach to and understanding of the legal order be accepted, several obstacles for legal AI and the automation of jurisprudence emerge. Law, under this philosophy, is not a closed system, nor can it ever be. There are many intrinsic limitations, gaps and imperfections. This contrasts with most ongoing attempts in legal AI that see or treat the legal system as closed and coherent. In Fuller's approach, law is also not a top-down project from either a heavenly legislator or an ultimate sovereign, but a codified form of customs. To understand the code, an understanding of the underlying customs and societal order is necessary. If legal AI truly tried to do justice to these foundational elements, arguably, this would defeat its purpose because this requires a decision tree that is so nuanced, detailed and granular that it imitates life: the map becomes as big as the territory. Finally, law is not unambiguous and factual nor can it ever be. It is a symbolic order that purposely creates and is dependent on legal fictions. This means not only that AI should be trained to bring about fictions and understand them, but also that because law is essentially symbolic, both in its essence and in its consequences, AI should be trained on symbolic meaning and effects.

## 3. The Pre-Conditions of Law

The previous section explained that law is not a closed system, but that it is based on implicit laws of lawmaking, customs and symbolic abstractions, and that it creates a new set of abstractions and social customs. This section explains what this means for the limits of the legal order by homing in on three particular aspects, namely the prerequisites for an effective legal order (Section 3.1), the prerequisites for a legitimate legal order (Section 3.2) and the prerequisites for attaining the intrinsic end of the legal order (Section 3.3). Finally, a small wind-up is provided (Section 3.4).

### 3.1 The Effectiveness of the Legal Order

In an essay on tyranny, Fuller describes the hypothetical story of a highwayman who gives people the choice between their money or their lives. 'Why does the highwayman not take people's money and then kill them?' Fuller asks rhetorically, only to point out that if the highwayman would do this, this would become known in the area in which he operates, and people would become scared, avoidant and, if they were to meet the highwayman, feel they have nothing to lose and put up a fight. The highwayman has an interest in being known as trustworthy. There is a reciprocal relationship between the robber and the robbed and, like in a pre-modern society, there is a complex web of interdependencies. The robber may not only face consequences from the robbed if he

does not keep his word, he may also face them from his fellow robbers (who may take action if he operates in a way that does harm to their reputation of 'honest thieves'), the relatives of the robbed may seek revenge, public outcry may lead to more police patrols, and so forth.

The same applies to a hypothetical 'rational tyrant'. Suppose there is a tyrant and, being a tyrant, their main aspiration is for everyone to obey them. The tyrant will discover that citizens will obey them more effectively if they are happy and satisfied with the role they play in the system. However, to be happy and satisfied, as a minimum, citizens need to feel that they are not merely means to the tyrant's ends; they must be approached as though they were ends in themselves. The tyrant will understand that citizens need not only develop capacities and skills that directly relate to their immediate assigned tasks, but they must also be able to fully explore their capacities and flourish as human beings in order to be happy and develop a broad set of capabilities, therewith maximising their 'usefulness' for the tyrant in the long run.[16]

Because language and laws are inherently ambiguous, the tyrant also needs citizens to have the mental capacity to understand the purpose of the law and the individual autonomy according to their own discretion.

> With this consideration in mind, our intelligent tyrant now proceeds to plan how to employ his subjects as tools for the realization of his purposes. A little reflection will remind him that he cannot effectively use another human being as a tool without according to him some power of choice, some opportunity to use his own discretion. When I hire the neighbor's boy to mow my lawn I do not begin by imposing on him a long and abstruse definition of what I mean by "lawn"; I assume he will have the good sense not to push the mower into my tulip bed just because he sees a few blades of grass growing up among the tulips. Simply from the standpoint of engineering efficiency in achieving a goal, some discretion and choice must, then, be accorded the human agent. This conclusion is reinforced when we recall that a favored and often successful mode of revolt is to carry out instructions with a wooden literalness; many a domineering parent has had his inclinations toward tyranny curbed by the retort, "But I did just what you told me to do!"[17]

16 This example is repeated by Fuller with regard to the gunman situation in: Fuller, L. L. (1964). Irrigation and tyranny. *Stan. L. Rev.*, *17*, 1021., p. 1027.

17 Fuller, L. L. (1968). Freedom as a Problem of Allocating Choice. *Proceedings of the American Philosophical Society*, *112* (2), 101-106, p. 105-106. A similar story is by Fuller when a babysitter is asked to teach the children a game and she teaches them to throw dice for money or to duel with a kitchen knife. Fuller, L. L. (1969) 'The Morality of Law', Yale University Press, London, p. 138.

There are three important lessons here. First, there is a constant reciprocal relationship between ruler and ruled: to rule most effectively, the ruler should allow the maximum freedom to the ruled and treat them as an end rather than a means. Second, even if there are no legal limits to what a tyrant can do, there are practical limits, which contrasts again with positive law theory.

> [I]n the literature of jurisprudence, law is generally defined as consisting of those rules that emanate from some human source that is itself regarded as formally authorized to enact or declare law. In the absence of explicit constitutional limitations, this human source can enact anything it sees fit into law. Its laws may be wise or foolish, intelligible or obscure, just or unjust, prospective or retrospective in effect, general or specific in their coverage, published or unpublished, etc.
>
> In all this variety, it is assumed there is no structural constancy, except that imposed by the formal rule which identifies the authorised source of law. But this view overlooks the fact that there are what may be called informal limitations implicit in any attempt to subject human conduct to the control of general rules.[18]

Third, the 'rational' tyrant contrasts with the irrational tyrant. An irrational tyrant has the same desire as the rational tyrant but attempts to achieve this desire in an irrational way. Thus, it is important to acknowledge that of course, law can be and is abused by tyrants all over the world. Although these tyrants may be effective in the short term, just like a highway robber who steals and murders, ultimately, such regimes are not stable and will come to an end. The instrument of law, by its very nature, by the very fact that it uses language which requires interpretation, by the very fact that humans need to understand the rules and implement these themselves, is opposed to suppression and requires human autonomy. The Spanish Inquisition could have chosen to torture people by putting them in a comfortable chair, but this would have been irrational because they would have been using the wrong instrument (the comfortable chair) in light of the goal to be achieved (torture). Similarly, it is irrational to believe that the use of law can result in obedient, numb citizens; tyrants use other instruments to that end, such as fear and violence, and use law as a pretext only.[19] The effectiveness of the legal order thus depends on human autonomy.

---

18 Fuller, L. L. (1964). Irrigation and tyranny. *Stan. L. Rev.*, 17, 1021.

19 This is precisely why he questioned whether the Nazi laws were laws proper: they were used as a pretext only. 'The German people were notoriously deferential toward authority, but even for them, as Hitler shrewdly saw, the habit of law observance was not a blind conditioned reaction toward orders coming from above, but was associated with a faith in certain fundamental processes of government, in particular with adjudication by disinterested judges and with statutes

### 3.2 The Legitimacy of the Legal Order

Supposing a rational ruler, there are eight routes for legislative failure, according to Fuller.

> The first and most obvious lies in a failure to achieve rules at all, so that every issue must be decided on an ad hoc basis. The other routes are: (2) a failure to publicize, or at least to make available to the affected party, the rules he is expected to observe; (3) the abuse of retroactive legislation, which not only cannot itself guide action, but undercuts the integrity of rules prospective in effect, since it puts them under the threat of retrospective change; (4) a failure to make rules understandable; (5) the enactment of contradictory rules or (6) rules that require conduct beyond the powers of the affected party; (7) introducing such frequent changes in the rules that the subject cannot orient his action by them; and, finally, (8) a failure of congruence between the rules as announced and their actual administration.[20]

These principles are grounded in the reciprocity between citizen and state. Citizens are generally held to obey and respect the rules, and with regard to most of the rules, follow them at least most of the time. Each of the eight routes for failure are in essence a violation of human autonomy. Indeed, every departure from these principles is:

> an affront to man's dignity as a responsible agent. To judge his actions by unpublished or retrospective laws, or to order him to do an act that is impossible, is to convey to him your indifference to his powers of self-determination. Conversely, when the view is accepted that man is incapable of responsible action, legal morality loses its reason for being.[21]

emerging from deliberative procedures participated in by elected representatives. Hitler therefore strove to preserve as a screen for his manipulations the familiar outer appearance of due process, realizing that law observance (and consequently the practical efficacy of his regime) was dependent on keeping that appearance as close as possible to reality. Where he could without sacrifice, he conformed to the demands of legality. But he had the advantage of being able to accomplish his will through either of two instruments, the state and the party. Where the result he desired would be embarrassed or compromised by those restraints that even the appearance of due process entails, he acted, not through the state, but through the party "in the streets."' Fuller, L. L. (1953). American Legal Philosophy at Mid-Century--A Review of Edwin W. Patterson's Jurisprudence, Men and Ideas of the Law. *J. Legal Educ.*, *6*, 457, p. 466. See also: Rajah, J. (2012). *Authoritarian rule of law: Legislation, discourse and legitimacy in Singapore*. Cambridge University Press.

20 Fuller, L. L. (1969). 'The Morality of Law', Yale University Press, London, p. 39.

21 Ibid, p. 162-163.

If citizens do not know the rules, are unable to adapt their behaviour to them or do not understand the rules, they cannot follow them. This means that the legal order is neither effective nor legitimate. For example, not only is it illegitimate to punish people for behaviour that was not unlawful at the time the conduct took place, but citizens can also not follow the rules retroactively. If the state structurally ignores these eight principles, in turn, citizens are not held to obey the laws or only to a limited extent.[22] There is thus an intrinsic relationship between the factual (can citizens obey the law; does the law abide by the rule of law) and the normative (should citizens obey the law; is this a valid law). These principles are consequently not pre-legal moral norms, such as natural rights, but what Fuller calls the 'inner morality' of law.

Although Fuller does not subscribe to any form of natural law in the classic sense, there might be a technological natural law.[23] The principles of 'inner morality' are in themselves neutral to the societal goals which legal orders aim at.[24] Thus, in principle, a Nazi-like regime could aim at morally-corrupted goals and respect the principles of legality at the same time. Still, there is a close affinity between legality and justice; unsurprisingly, morally-corrupted regimes are inclined to ignore the rules of legality.[25] Even the most corrupted regimes are afraid to explicitly and openly acknowledge the horrors they commit or want to commit. Often, the worst atrocities are based on ad hoc decisions without any explicit legal basis and without those decisions being communicated openly to the public.[26] Thus, they publicly pretend to uphold the principles of the rule of law; a normal legal regime is often in place, while atrocities are based on extra-legal instruments.[27]

However, there is an additional layer of complexity. It is not true that any violation of one of the eight principles will per se lead to a law being null and void, as sometimes such violation is a necessary evil. The principles of the rule of law are often understood as minimum principles of law, which they are, but they are also aspirations. Only in a utopia can all eight elements be respected to a full extent, since some situations will always necessitate deviation from those principles. For example, there was no law in

22 Ibid, p. 39.

23 Fuller, L.L. Means and Ends, p. 63. In: Fuller, L. L. (1981). 'The Principles of Social Order', Duke University Press, Durham.

24 Fuller, L.L. (1969). 'The Morality of Law', Yale University Press, London, p. 152.

25 Fuller, L.L. (1958). 'Positivism and Fidelity to Law: A Reply to Professor Hart', 71 Harvard Law Review 630, p. 631.

26 Also, Fuller believed "that coherence and goodness have more affinity than coherence and evil." Fuller, L.L. (1957). 'Positivism and Fidelity to Law: A Reply to Professor Hart', 71 Harvard Law Review 630, p. 636.

27 See further: J. Rajah, Authoritarian rule of law: Legislation, discourse and legitimacy in Singapore, Cambridge University Press, 2012. Naím, M. (2022). *The revenge of power: how autocrats are reinventing politics for the 21st Century*. St. Martin's Press.

place during the Nazi regime that would prohibit certain acts of terror, murder and annihilation, but it is clear that at least the Nazi-leaders had to be tried after WWII, by retroactively applying legal standards if necessary. Moreover, the eight principles of law conflict with each other. For example, some communist countries endeavoured to make laws so clear that they would be intelligible even to the illiterate working class.[28] This, however, came at the cost of legal consistency and the overall coherence of the legal system. If one of the eight principles cannot be upheld, this can be remedied by strong performance on another of the eight principles. They are interdependent principles of legality geared towards the same end.[29]

### 3.3 The Teleology of the Legal Order

A teleological perspective implies that a creature inherently strives towards an intrinsic end. A tree, for example, naturally strives to catch sunlight and water to grow. Thus, it will grow roots in the ground and branches and leaves. A good and healthy specimen is a tree that has the instruments to achieve its innate goal; a tree that grows no leaves and has few roots is unhealthy, weak and unable to fulfil its purpose. A teleological perspective can also be adopted with respect to manmade objects, which have been designed with a purpose in mind. As such, a law may be compared to a chair, and the art of law-making to the craft of carpentry.[30] There are both minimum limits and an inherent aspiration on which a manmade tool must be judged. A chair must have legs and a seating to be called a chair. If a chair has one uneven leg, we might call it dysfunctional, if it lacks a leg altogether, we might call it defect or broken, and if it has no legs whatsoever, we might call it a cushion instead of a chair. As an aspirational quality, a chair may be designed to provide the user with the most comfortable experience possible, or it must help the user sit in the most ergonomic way possible.

Fuller speaks of two moralities in this regard: the morality of duty and the morality of aspiration. In terms of language, the distinction can be compared to the difference between the rules of grammar and those of aesthetics.[31] The difference can also be linked to the difference between life, with the state as an instrument to provide for the basic necessities of life, and the good life, with the state as a facilitator for people to flourish as human beings. To be effective and legitimate, law must not undermine the autonomy of citizens and, as an aspiration, it must enable maximum autonomy and freedom for citizens. Law is an instrument that is intrinsically geared towards maximising human freedom.

28 Fuller, L.L. (1969), 'The Morality of Law', Yale University Press, London, p. 93.
29 Ibid, p. 104.
30 Ibid, p. 96.
31 Ibid, p. 18.

> The only permissible form of legislation is the sort that lets individuals plan their own lives. Simply put, legislative enactments are baselines for self-directed conduct by citizens, providing the minimal restraints necessary for continuing interaction. Legislation properly conceived permits citizens to order their own affairs, to pursue their own good in their own way (in the words of John Stuart Mill).[32]

### 3.4 Conclusion

The previous points have several implications for legal AI and the automation of jurisprudence. Both the effectiveness and legitimacy of the legal order ultimately lay in granting citizens autonomy in interpretating the meaning of rules and allowing them to assess how the rules should be interpreted. Not only are these preconditions for the legal order; optimising freedom and autonomy are its ultimate aspiration. In essence, this would mean that legal AI should assess whether its operational quality is that of furthering human autonomy as a whole. To make matters more complex, the principles that drive and protect human autonomy, the principles of the rule of law, are not only minimum conductions for legal orders, but they are also aspirations that can never be attained in full, if only because attaining one ideal comes at the cost of attaining the others.

## 4. Law As an Internal Interdependency Between Opposition Elements

The legal order is essentially a deep interplay between different yet interdependent elements. It is the interplay between fact and fiction, between aspiration and realisation, between order and freedom, between means and ends. This means that codification of the legal order should, following Fuller's theory, take account of these interdependencies. This section provides three examples of what type of interplays should be taken into account. First, the interplay between the rule of law and democracy (Section 4.1). Second, the interplay between order and freedom and between negative and positive freedom (Section 4.2). Third, like the pre-modern society is dependent on a complex web of reciprocal relationships between non-formal parties, the legal order is dependent on a complex web of reciprocal relationships between formal actors (Section 4.3). Finally, a small wrap-up is provided (Section 4.4).

32 Editor's note, Fuller, L.L. 'The implicit Laws of Lawmaker', p. 158. In: Fuller, L.L. (1981). 'The Principles of Social Order', Duke University Press, Durham, 1981. See also: Fuller, L. L. (1964). Irrigation and tyranny. *Stan. L. Rev.*, *17*, 1021.

### 4.1 Democracy and the Rule of Law

Most human associations start from a shared goal or need. Food, shelter and security may have been typical goals for coming together in pre-modern societies, but most associations are still based on a shared goal. Fuller gives the example of a book club he and his friends started when they were young and recounts a classmate who wanted to join. After some deliberations, the members decided he could join but, when he did, it soon became clear that his level of knowledge was insufficient. He was expelled without a process, without him being heard, and without any formal reason being provided for his expulsion.

The bigger the organisation and the higher the stakes of being admitted or expelled, the more need there is for rules, procedures and due process, which is why big private institutions (e.g. the Church) have elaborate internal codes. Processes, functions and entitlements are formalised and codified. Decisions about admission and expulsion are not based on sentiments but on rules and formalities. Thus, two ideal types of organising can be defined along a spectrum. On one end of this spectrum is the book club that is dominated by a shared commitment and is largely unformalized. On the other end of the spectrum are associations dominated by rules and procedures.

**TABLE 1.** *Two ideal types of organisations and their characteristics*

| Associations based on a shared commitment | Associations based on legal principles |
|---|---|
| Subjective | Objective |
| Group | Aggregate of individuals |
| Democracy | Rule of law |
| Friend-foe decisions | Due process |
| Substantive justice | Procedural justice |

No organisation exists that solely belongs to one category. Groups that function on the basis of reciprocal relationships and customs give rise to customary law, which will later be codified. Legal systems need to have a goal or commitment, as rules serve a certain purpose and aim towards something. One problem, however, is that associations tend to lean increasingly towards legality, bureaucracy and procedural justice. Fuller distinguishes between eight laws of human association or natural processes that occur when an association is set up, the final three of which are worth quoting here:

> Sixth law. In the normal course of its development an association tends to move toward dominance by the legal principle. At the end of this development, the principle of shared commitment commonly sinks into a state of quiescence until some

> crisis, such as an external threat to the association, brings it back to life, commonly in an awkward kind of rebirth and often in a new form. The tendency of the element of commitment to sink out of sight applies not only to what may be called the substantive commitments of the association, but also … to the prosaic kind of commitment that is an indispensable support of the legal principle itself. Seventh law. Once under way the development toward dominance by the legal principle feeds on itself and becomes accelerative. The aging association commonly displays the symptoms of what may be called creeping legalism. Eight law. The conditions of modern institutional life tend strongly to break down the distinction between the law of the political state and the internal law of associations. The result is an expansion of the jurisdiction off the regular courts of law to pass in review decisions of associations.[33]

For Fuller, neither one of the two principles of association should dominate, as that would undermine its effectiveness and legitimacy.

### 4.2 Eunomics

Means and ends are interdependent elements of the legal order. Forms liberate[34] as much as liberty forms. The end dictates the means as much as the means have an impact on which ends can be pursued. Suppose one was devising a not yet existing game and one would propose to first decide on what kind of enjoyment the projected game should serve and only afterwards decide on the rules of the game itself – it would be quite difficult.

> The pleasures derived from any form of play are always complex, and the enjoyment yielded by any particular game is the unique product of its own peculiar nature. If we are to invent a game, we shall have to start with ends vaguely perceived and held in suspension while we explore the problem of devising a workable system of play.[35]

Deciding on the ends and means is a heuristic and iterative process.

Law, in this respect, is similar to language. Although language is often unsuitable to precisely say what we mean, the limitations are necessary if we want to communicate

33 Fuller, L. L. Two Principles of Human Association, p. 92. In: Fuller, L. L. (1981). 'The Principles of Social Order', Duke University Press, Durham..

34 "[T]o become effective, freedom requires a congenial environment of rules and decisions." Fuller, L. L. (1954). Freedom--A Suggested Analysis. *Harv. L. Rev.*, *68*, 1305. See in detail: Rundle, K. (2012). *Forms liberate: reclaiming the jurisprudence of Lon L Fuller*. Bloomsbury Publishing.

35 Fuller, L. L. (1953). American Legal Philosophy at Mid-Century--A Review of Edwin W. Patterson's Jurisprudence, Men and Ideas of the Law. *J. Legal Educ.*, *6*, 457, p. 479.

at all.[36] In addition, certain languages are more apt, for example, for sharing emotions, others more for factual descriptions. Likewise, Fuller stresses, freedom as a form of social ordering, and order as a means to attain freedom, are interrelated:

> Society is, of course, impossible without some limits on individual freedom, that is, without some form of constraint. If freedom means the absence of constraint, the problem then becomes that of avoiding constraints at particular points and for particular reasons – or, reversing the emphasis, of assigning sound reasons for imposing constraint at particular points. The respective functions of freedom and constraints are, therefore, two aspects of the same question. The pattern of freedom is the reversed image of the pattern of constraint. The two form the structure of society as a whole.[37]

The legal order is an equilibrium between interrelated elements, such as means and ends, form and content, order and freedom. Similarly, the legal order should promote both positive and negative freedom, the one cannot do without the other. For example, the freedom:

> to cast one's ballot loses its point where the voter is not also free from restraints that prevent him from voting as he decides he should. The ridiculous "elections" held under the Nazi regime are an extreme case in point. A "freedom from" certain kinds of interference must, then, be presupposed in every "freedom to."[38]

For Fuller, eunomics (a concept inspired by the virtue ethical term *eudaimonia*, the good life), or good lawmaking, is the capacity of lawmakers to account for the interrelation between opposing elements, such as democracy and the rule of law, means and ends, negative and positive freedom.[39] These are not pairs of ideals that conflict with each other and should be kept in balance; they are interdependent. Democracy can only be achieved through the rule of law, just as the rule of law is geared towards human autonomy; positive freedom can only be achieved through negative freedom, just like how the essence of negative freedom serves as a facilitator for positive freedom; liberty can only be achieved through order, just like how order liberates.[40]

---

36 Fuller, L. L. (1968). Freedom as a Problem of Allocating Choice. *Proceedings of the American Philosophical Society*, *112* (2), 101-106, p. 102.

37 Fuller, L. L. 'Means and Ends', p. 59. In: Fuller, L. L. (1981). 'The Principles of Social Order', Duke University Press, Durham.

38 Fuller, L. L. (1954). Freedom--A Suggested Analysis. *Harv. L. Rev.*, *68*, 1305.

39 Fuller, L. L. (1953). American Legal Philosophy at Mid-Century--A Review of Edwin W. Patterson's Jurisprudence, Men and Ideas of the Law. *J. Legal Educ.*, *6*, 457,p. 475-476. See also: Fuller, L. L. (1958). Human purpose and natural law. *Nat. LF*, *3*, 68.

40 It is a philosophical understanding of democracy and the rule of law that Habermas would later

## 4.3 Reciprocity of Relationships

For Fuller, every branch of law is based on and facilitates reciprocity.[41] To the extent that laws are not aimed at facilitating human interaction, these tend to be amongst the most ineffective and controversial laws, such as the criminalisation of behaviour that has no victims (e.g. the use of marijuana, homosexual practices or gambling). Customs and customary law, as explained, are based on reciprocity and relational expectancies. Contract law, like customary law, is grounded in customs and geared towards reciprocal relationships.[42] International law, Fuller points out, while partially codified, is mostly a matter of interstate reciprocity and agreements, just like how administrative law is based on facilitating human interaction. Traffic rules, for example, are meant to guide and facilitate human interaction, rather than to correct behaviour. Even criminal law is best explained in terms of reciprocity. The prohibition of murder, for example, not only prohibits the type of action that violates someone's basic sense of autonomy and dignity, it also breaks the vicious circle of eye-for-an-eye, in which interhuman relations are annulled instead of promoted.[43] To Fuller, consequently, criminal law should not be seen as an instrument of social control but as a way to facilitate human interaction.[44]

> [A] rule against murder, effectively enforced, serves to enlarge the scope of the individual's interactions with others. In many of our cities are areas that strangers cannot enter without some risk to their physical safety. Here a failure of legal control results in a restriction on interaction, an interaction that in the long run might promote reciprocal understanding and, with it, a reduction in the risks that now aggravate distrust.[45]

In addition, the legal order consists of a complex web of interdependent reciprocal relationships between the various actors, such as the legislative, executive and judicial

qualify as the internal relationship between the rule of law and democracy. Habermas, J. (1995). On the internal relation between the rule of law and democracy. *European Journal of Philosophy*, *3* (1), 12-20.

41 Vise versa, customary law always contains, at least in embryonic form, elements statutory law or legal procedures.

42 Fuller, L. L. (1969). Human interaction and the law. *Am. J. Juris.*, *14*, 1p. 224. Fuller, L. L. (1941). Consideration and form. *Columbia Law Review*, *41* (5), 799-824, p. 806.

43 Fuller, L. L. (1969). Human interaction and the law. *Am. J. Juris.*, *14*, 1, p. 231-232.

44 Fuller, L. L. (1975). Law as an instrument of social control and law as a facilitation of human interaction. *BYU L. Rev.*, 89, p. 90.

45 Fuller, L. L. (1975). Law as an instrument of social control and law as a facilitation of human interaction. *BYU L. Rev.*, 89, p. 90.

branch.[46] A first example follows from the discussion of the rational tyrant: the reciprocal relationship between the ruler and the ruled.

> On the one hand, the lawgiver must be able to anticipate that the citizenry as a whole will accept as law and generally observe the body of rules he has promulgated. On the other hand, the legal subject must be able to anticipate that government will itself abide by its own declared rules when it comes to judge his actions, as in deciding, for example, whether he has committed a crime or claims property under a valid deed. A gross failure in the realization of either of these anticipations – of government toward citizen and of citizen toward government – can have the result that the most carefully drafted code will fail to become a functioning system of law.[47]

Autonomy and discretion are also important for both the executive and the judicial branch when interpreting and applying laws.

> Before the case is brought to court the defendant has to be arrested, and it would certainly be a rare policeman who routinely – and without taking into account the nature and circumstances of the offense – arrested every person he believed to have committed a crime. Certainly in dealing with minor offenses the police officer uses, and is expected to use, "judgement"; this judgement is inevitably affected by his perception of the kind of person the suspected party seems to be. When the case is brought to the prosecutor he in turn is influenced in some degree by similar considerations in deciding whether to prefer charges.... If the case goes to trial and the accused is found guilty, the question of the appropriate sentence has to be decided. In deciding that question the judge will take into account what is known about the defendant himself, his past, and his probable future propensities. Similar considerations will, of course, determine the granting of parole or a pardon.[48]

The judiciary and the legislative branch in particular have a reciprocal relationship.[49] A judge must try to understand what the legislative branch aimed at when adopting a law. If courts are too literal and take too textual an approach when interpreting a law, the legislative branch will take that into account when drafting laws. By contrast, if courts

46 Fuller sees the 'lawyer as an architect of social structures'. Fuller, L.L. The lawyer as an architect of social structures, p. 286. In: Fuller, L. L. (1981). 'The Principles of Social Order', Duke University Press, Durham.

47 Fuller, L. L. (1969). Human interaction and the law. *Am. J. Juris.*, *14*, 1.

48 Fuller, L. L. (1969). Human interaction and the law. *Am. J. Juris.*, *14*, 1.

49 Fuller, L.L. The Role of Contract, p. 195. In: Fuller, L. L. (1981), 'The Principles of Social Order', Duke University Press, Durham.

suddenly adopt a significantly more liberal approach than was intended, this will likely lead to legislative action.[50] Furthermore, there is an interdependent relationship between courts. Courts have to take account of each other's judgments and cannot ignore or habitually deviate from standing jurisprudence. At the same time, each court must use its own discretion and assess each matter in light of the specific circumstances of the case.[51] Finally, a court hearing essentially facilitates the discussion between the two opposing parties under streamlined conditions and formalised procedures.[52] A judge must hear all parties in a case and allow them to respond to questions and remarks; the judgment is the end-result of human interaction.

### 4.4 Conclusion

For legal AI, the findings from this section mean that there are no clear boundaries or binary choices but a complex web of interdependencies. AI, to be effective, should take account of each and every one of these interdependencies and codify them. This would also require the codification of a heuristic and iterative process between means and ends, between freedom and order and between democracy and the rule of law.

## 5. Case Law

From Fuller's perspective, there are important limitations to case law that are relevant when building AI systems that would deliver judgments. First, many legal disputes are best addressed outside the legal system (Section 5.1). Second, one of the essential roles of judges is to interpret language and meaning about which there is no certainty, only ambiguity (Section 5.2). Third, there is often no 'good', 'right' or 'best' solution, but several opposing yet equally appealing potential outcomes, not only in terms of interpreting what a law means, but also for arriving at an outcome (Section 5.3). This section ends with a small conclusion (Section 5.4).

50 Fuller, L. L. 'Some presuppositions shaping the concept of "Socialization"', p. 39-40. In: Tapp, J. L. & Levinne, F. J. (1977). 'Law, justice and the individual in society: psychological and legal issues', New York, Holt, Rinehart and Winston

51 Fuller, L. L. (1975). 'Law as an Instrument of Social Control and Law as a Facilitation of Human Interaction', Birmingham Young University Law Review 89, p. 95.

52 Fuller, L. L. Forms and Limits of Adjudication, p. 121. In: Fuller, L. L. (1981), 'The Principles of Social Order', Duke University Press, Durham.

## 5.1 Mediation

Fuller was one of the early advocates of mediation. The fundamental role of the mediator is to reorient parties toward each other,

> not by imposing rules on them, but by helping them to achieve a new and shared perception of their relationship, a perception that will redirect their attitudes and dispositions toward one another. This suggest[s] a certain antithesis between mediational processes, on the one hand, and the standard procedures of law, on the other, for surely central to the very notion of law is the concept of rules.[53]

A judicial approach is ill-suited, he suggests, when the relationships between parties are complex, when multiple parties are involved and/or when there is a high number of interdependencies. These situations resemble pre-modern societies.

> I am contending that when we encounter social contexts similar to those of a primitive society, we too resort to mediative rather than adjudicative methods of problem solving. (I would prefer "problem solving" to the fashionable "dispute resolution"). The contexts that make mediation preferable are in general those of heavy and complex interdependencies. Among such contexts in our society one might list: (1) marriage, (2) closely held corporations, (3) tenants packed together in public housing, (4) co-authors of a book. In all these cases some "straightening out" may be essential, rather than an adjudicative determination of rights and duties.[54]

The attribution of custody over children after divorce, for example, is least served by a judge making a call because both parties cannot come to an agreement. The best result is that parents remain on speaking terms, go to school meetings together, make important calls concerning their children together and know how to reach each other if an emergency arises. This cannot be achieved through a judicial procedure, which generally deepens the conflict. Mediation is an intra-extra-legal element, like private associations that set their own rules, procedures and customary law (like the Church in Western democracy). It is not outside the legal order, as the legal regime facilitates and sometimes even encourages mediation, but it is an essentially non-juridical way of dispute resolution. Having conflicting parties talk again about their shared interests and understandings is something that is typically difficult to achieve through legal automation.

---

53 Fuller, L.L. (1970). Mediation--its forms and functions, Southern California Law Review, *44*, 305, 1070.

54 Ibid.

## 5.2 Judicial Interpretation of Language Meaning

When interpreting a law, the judge needs to consider the intentions of the lawmaker, the circumstances under which the law was adopted and the prevailing customs at that time. Fuller underlines the 'direction-giving quality of purposive facts'. If a mechanic with poor English skills were to write an instruction on how to build a machine and two persons, an English professor and another mechanic, were to read their instructions, then the latter would not get lost in the 'literal or factual' interpretation of the text but try to find its essence and thus understand the instructions best, Fuller poses. Facts and literal meaning cannot be understood without their value and purpose.

> As for the application of the dichotomy of is and ought to the law, it is fairly clear that with legal precepts, as with the instructions for assembling a machine, what a direction is can be understood only by seeing toward what end result it is aimed. The essential meaning of a legal rule lies in a purpose, or more commonly, in a congeries of purposes. Within the framework of this purpose, or set of related purposes, the sharp dichotomy between fact and evaluation cannot be maintained; the "fact" involved is not a static datum but something that reaches toward an objective and that can be understood only in terms of that reaching.[55]

This has important repercussions. Imagine there was a rule that prohibits people from sleeping at the train station, which was adopted because there were multiple homeless persons sleeping on the benches at the train station day and night. Suppose Person A's connecting train is delayed for hours and, sitting on the bench at night waiting for the delayed train to arrive, they doze off for a few seconds. Person B, to the contrary, does not sit on the bench but lays down on it, using it as their mattress but making sure to stay awake. If both A and B were to be brought before court, a literal interpretation of the rule would lead a judge concluding that the first person was in violation of the law and the second was not, while the spirit of the law would point to the exact opposite interpretation. Fuller would suggest the latter is the best interpretation, which means that legal AI should interpret rules according to their meaning, rather than their literal text.

It is not only the purpose of the law that needs judicial interpretation. Since the legal order is essentially a symbolic order built on concepts and abstractions, interpretation is also required when it comes to basic concepts used in a law. Suppose there was a law prohibiting 'vehicles' in the park. This would require a determination of what counts as

55 Fuller, L. L. (1953). American Legal Philosophy at Mid-Century--A Review of Edwin W. Patterson's Jurisprudence, Men and Ideas of the Law. *J. Legal Educ.*, *6*, 457, p. 470-471.

a vehicle and what does not. Fuller suggests that there are core examples that clearly fall within this term, but that there is also a penumbra of ancillary concepts that may or may not fall under the definition, depending on the circumstances. Ought the law, for example, be interpreted as barring children's tricycles, buggies and an old military tank placed on a pedestal in memory of the victims of WWII?[56] Answers may be partially found in the explanatory memorandum of the bill, but it is impossible for a lawmaker to discuss all potential objects. In addition, a judge would need to make a decision on the extent to which new vehicles that emerged after the adoption of the law fall under the scope of the law and the intention of the lawgiver.

### 5.3 The Case of the Speluncean Explorers

Consider a hypothetical case in which a group of spelunkers get stuck in a cave for several days, and they decide to kill and eat one of their number to survive. The question that a court would then have to decide on would be whether the surviving explorers should be convicted of murder or not. Fuller presents five separate reasonings by the five judges sitting on this case, each having their own take on the question. Each of the judges uses valid legal doctrines and sound legal reasoning. Still, they arrive at different conclusions.

- Judge 1 stresses that the law is clear: it is murder. The explorers consciously killed a person, not out of self-defence. Judges must, Judge 1 argues, apply the law, not make it. That is for the legislative branch. Emphasising the importance of the separation of powers, they suggest the executive branch pardons the survivors (which is their prerogative, not that of the judges). They point to the extreme situation in which the defendants found themselves.

  > [I] think we may therefore assume that some form of clemency will be extended to these defendants. If this is done, then justice will be accomplished without impairing either the letter or spirit of our statutes and without offering any encouragement for the disregard of law.[57]

- Judge 2 states that the survivors found themselves in an exceptional situation and that the legal order provides the rule, not the exception. *Necessitas non habet legem*, the legal regime does not apply to these extreme scenarios. Moreover, one of the aims

56 Fuller, L. L. (1957). Positivism and fidelity to law--A reply to Professor Hart. *Harv. L. Rev.*, *71*, 630, p. 662.

57 Fuller, L. L. (1948). The case of the speluncean explorers. *Harv. L. Rev.*, *62*, 616.

of criminal law is to have a deterrent effect. Judge 2 points out that would the court convict the defendants for murder, there would be no such effect because of the exceptionality of the case.

> When a situation arises in which the coexistence of men becomes impossible, then a condition that underlies all of our precedents and statutes has ceased to exist. When that condition disappears, then it is my opinion that the force of our positive law disappears with it. We are not accustomed to applying the maxim *cessante ratione legis, cessat et ipsa lex* to the whole of our enacted law, but I believe that this is a case where the maxim should be so applied.[58]

- Judge 3 states that criminal law serves several goals, such as retribution and rehabilitation. The various objectives underlying the law cannot be reconciled in this case. That is why Judge 3 does not arrive at a verdict.

> Since I have been wholly unable to resolve the doubts that beset me about the law of this case, I am with regret announcing a step that is, I believe, unprecedented in the history of this tribunal. I declare my withdrawal from the decision of this case.[59]

- Like Judge 1, Judge 4 finds that the law is clear: murder. However, unlike Judge 1, Judge 4 thinks that advising the executive branch on the appropriate course of action would go against the separation of powers. It is also wrong for a judge to warn against the consequences of democratically adopted laws. If a democratic legislator finds that this effect of the law it has adopted is undesirable, it should change or revise the law. These potentially undesirable effects can at times only become clear if judges show what a law means in practice; smoothing out negative or undesirable effects in this sense undermines the potential for a democracy to correct itself.

> I believe that judicial dispensation does more harm in the long run than hard decisions. Hard cases may even have a certain moral value by bringing home to the people their own responsibilities toward the law that is ultimately their creation, and by reminding them that there is no principle of personal grace that can relieve the mistakes of their representatives.[60]

58 Ibid, p. 1854.
59 Ibid, p. 1863.
60 Ibid, p. 1868.

- Finally, Judge 5 refers to public opinion and common sense. Both factors lead this Judge to the conclusion that the death penalty should not be imposed, and the conviction set aside.

> Now I know that my brothers will be horrified by my suggestion that this Court should take account of public opinion. They will tell you that public opinion is emotional and capricious, that it is based on half-truths and listens to witnesses who are not subject to cross-examination. They will tell you that the law surrounds the trial of a case like this with elaborate safeguards, designed to ensure that the truth will be known and that every rational consideration bearing on the issues of the case has been taken into account. They will warn you that all of these safeguards go for naught if a mass opinion formed outside this framework is allowed to have any influence on our decision.[61]

However, Judge 5 points inter alia to the jury system in place in legal systems around the world and holds that this is an explicit attempt to include the common or layman's view on matters in the legal system. The same may to an extent be said for the executive's prerogative to pardon a criminal, which is an inter-extra-legal power, through which the king or president usually considers public opinion.

It is important to Fuller that each of these positions are perfectly legitimate and go back to philosophical debates that date as far back as the days of Plato and Aristotle.[62] The added benefit of law is consequently not (only) that it gives correct answers to legal questions; rather, through the legal reasoning, the debate that precedes a verdict (with both parties bringing forth different arguments) and the explication of the verdict, a judgment begets its legitimacy.

## 5.4 Conclusion

The foregoing means that, for legal AI and the automation to be successful, several aspects should remain beyond its scope, namely those where the most preferable solution is not legal but relational in nature. Normally, judges make such assessments and may suggest parties start mediation processes; instead, AI would need to assess when parties should sit down together instead of delivering a verdict. This requires emotional comprehension. For example, is it still possible to have divorced parents work together in their child's best interest or are both so angry at each other that having them sit together would only lead to further escalation? In addition, law needs to be codified in

61 Ibid, p. 1870.
62 Ibid, p. 1874-1875.

such a way that it does justice to the ambiguity of language and, preferably, gives several equally appealing outcomes for legal disputes. This, in turn, would mean that legal AI cannot be used for automatic decision-making or for truly replacing judges but for providing judges with several potential outcomes at most. The legitimacy of a ruling, Fuller would hold, does not only depend on its content but also on the process of being heard as a party by a judge, seeing the judge assess all relevant aspects of the case, and having the judge read out their verdict in person.

## 6. Conclusion

This chapter briefly discussed the three most important theories in philosophy of law: legal positivism, natural law and legal pragmatism. It has suggested that many of the promises of legal AI align with the approach to and understanding of law as accepted by legal positivists. Legal positivists regard law as a closed system of hierarchically structured rules that should, ideally, be as unambiguous and consistent as possible. It rejects extra-legal elements such as customs, intentions and morals from legal interpretation. Instead, it focusses on the prerogative of the legislative branch and favours a textual interpretation of the law. The legal order, in its ideal form, comprises of a blueprint for a society-to-be; it is devoid of symbolism and legal fictions. If law is understood through the lens of natural law theory, legal AI becomes more problematic. There are some aspects that this philosophy of law shares with legal positivism, such as its top-down approach and its belief in a hierarchal set of clear, unambiguous and internally consistent rules. Other aspects, however, lend themselves less to legal automation, such as that it perceives law as a teleological instrument and that it suggests that human laws cannot derogate from natural rights, such as respect for human dignity.

If law is conceived from a legal pragmatist perspective, however, it becomes clear that the legal order cannot be automated, that is, if law were codified through AI, the essence of the legal regime would be lost. Not only is the legal order deeply dependent on extra-legal elements, such as language, customs and intentions, but extra-legal ways of dispute resolution are often preferable. There is often no right answer when interpreting legal texts and applying them to a concrete case. The legitimacy of a judgment is dependent not only on the correct interpretation of the law but on social factors, such as being seen and heard by another human being. Perhaps most importantly, the legal order is a living instrument dependent on the interdependency between the rule of law and democracy, between means and ends, between fact and fiction, between order and freedom. Lawmakers, judges and officials are in a constant interplay, just like citizens and the state.

**TABLE 2.** *Core differences between the three main branches of legal philosophy*

| Legal Positivism | Natural Law Theory | Legal Pragmatism |
|---|---|---|
| Top-down legal order | Top-down legal order | Bottom-up legal order |
| Legal order designed by humans | Legal order designed by God/Nature | Legal order arises organically from customs |
| Prerogative with the (human) legislative power | Prerogative with the (divine) legislative power | Reciprocity between governmental powers |
| Citizens obey the law; lawmaker unbound | Citizens obey the law, unless lawmaker violates minimum conditions | Reciprocity between citizen and state |
| Separation of is and ought | Minimum requirements | Is and ought inseparable |
| Textual approach, literal order | Both textual and grammatical approach | Grammatical approach, symbolic order |
| Closed system (the ideal of law is that all extra-legal elements are excluded) | Minimum norms (pre- and supra-legal norms) | Open system (the whole legal system is based on and facilitates extra-legal elements) |
| Neutral to content of laws | Minimally concerned with content of laws | Based on and geared at freedom and autonomy |
| Right answer in a concrete case exists (knowledge accessible to humans) | Right answer in a concrete case exists (knowledge accessible to God) | Multiple equally appealing and legally correct answers exist, especially in hard cases |

It is difficult to see how AI could codify such a complex system of interdependencies. It would require knowledge of the context in which laws where adopted and case law issued. It would require an understanding of words and concepts and how they have shifted over time. It would require knowledge of the intention of the lawmaker and the judge. It would require an understanding of customs. It would require an understanding of symbols, metaphors and fictions. It would require an understanding of the intrinsic limits of laws and determine when extra-legal forms of dispute resolution might be preferable. It would require an overall assessment of the equilibrium between means and ends, between negative and positive freedom, between the rule of law and democracy. It would require an evaluation of the balance between the various branches of state and the reciprocal relationship between citizens and the state. It would require an evaluation of whether the outcomes of laws and case-law ultimately further human freedom and autonomy and it should account for the fact that in hard cases, there are no right or wrong answers and that more in general, the legitimacy of the legal system does not depend on its rational and consistent application, but on human reciprocity.

Could legal AI not be used in the clearest cut and most highly repetitive cases? No, Fuller's approach would suggest, because there are no clear-cut cases. Legal rules should

be interpreted according to their meaning, not according to the text, and what purposive facts mean always requires interpretation. No, also because even if it is clear a matter that falls within the scope of a law, its application may still be counter to its intention, such as with a person who dozes off on a bench at the station while waiting for a delayed train. No, Fuller would suggest, because AI would need to assess whether this is a matter that should be dealt with in extra-legal ways, which requires an emotional understanding of the parties involved. No, Fuller would say, because there are often many correct legal approaches to a single case; legitimacy ultimately depends on a party to a case feeling they are being seen, heard and involved in the process that leads up to the ultimate verdict.

PART

# V

# REVOLUTIONISING EUROPEAN TECHNOLOGY REGULATION

CHAPTER

# XIII

# Conclusion

**Bart van der Sloot,**

**Giorgio Monti &**

**Friso Bostoen**[1]

https://doi.org/10.26116/gzdp-7x89

1 Associate, Full and Assistant Professor, Tilburg Institute of Law, Technology, and the Society (TILT), Tilburg Law School (TLS), Tilburg University.

## 1. Introduction

This book contains the first outcomes of a project that aims to revolutionise European technology regulation by shifting the focal point of the European legal acquis from personal behaviour to data. The project is still ongoing, which is why this book does not always present answers or detailed proposals. Many chapters map the problems with the current regulatory regimes and show why the underlying philosophies no longer work. Although some have indicated a possible direction ahead, each chapter has made it clear that alterative regulatory approaches are few and far between, and each come with their own difficulties. The main conclusions of each chapter are briefly recapitulated in section 2. Finally, section 3 maps some of the more important questions that this research project has yielded thus far, which will serve as a basis for future research. The hope is that these will not only be picked up by researchers from the Tilburg Institute for Law, Technology and Society, but that they will inspire researchers around the globe.

## 2. Recapitulation

Part II contained five shorter chapters that show how the data-driven environment challenges classic divides that underpin many current legal frameworks, in particular the divides between the private and public domains, the private and public sectors, and private and public law.

Chapter 2 discussed the case of Clearview AI to illustrate the increased reuse of publicly available personal data acquired via web scraping. Law enforcement agencies have relied on publicly available social media communications. Researchers use web scraping to collect and analyse large scale datasets, and the disclosure of personal data to the larger public may be mandated by statutory law in light of transparency obligations. The question thus arises of which privacy rights individuals have in the digital public space and how they can be enforced effectively. The problem is that there is no good or commonly accepted definition of the digital public space, nor is there a consensus on informational privacy rights in the digital public domain. The chapter recommended that future research focuses on reconceptualising both the digital public domain and the role of privacy in the data-driven environment.

Chapter 3 homed in on transatlantic personal data transfers between the EU and the US for Anti Money Loundering (AML) and Countering Financing of Terrorism (CFT) purposes. It showed that, even apart from the distinct privacy and data protection approaches, there are differences between the AML/CFT laws of the two jurisdictions. Both the EU and the US have adopted a risk-based approach to AML/CFT, but the EU approach is more prescriptive, with specific requirements for customer due diligence,

beneficial ownership, and the reporting of suspicious transactions. Comparatively, the US allows more flexibility for institutions to tailor their AML/CFT programs to their specific risk profiles. This creates uncertainty and confusion for legal entities, organisations and individuals. Consequently, this chapter found that there is an urgent need for ending the uncertainty and demonstrated that future research should aim to accurately indicate and demarcate as many aspects of the existing problems as possible, as well as provide regulatory recommendations to solve them.

Chapter 4 analysed the use of DNA data as evidence in criminal cases and the recent use of privately owned DNA databases by law enforcement authorities in particular. The European court of Human Rights (ECtHR) jurisprudence on this point is far from conclusive, providing only general frameworks. It is clear that the ECtHR does not consider DNA retention as such to go against the European Convention on Human Rights, but it does underline the importance of implementing sufficient legal and practical safeguards against the misuse of the collected material and data. The EU Law Enforcement Directive requires law enforcement authorities to have a clear legal basis regarding DNA retention that does not lead to blanket and indiscriminate collection of genetic samples and clearly prescribes the attached safeguards against data misuse and abuse. However, there are various ways to circumvent these legal restraints. This is why the chapter argued that clear regulatory frameworks should be developed to provide adequate safeguards.

Chapter 5 pointed to the increasing evidence that using algorithmic technologies in election campaigns can impact free elections. The most powerful techniques that political campaigners have been using extensively since 2015 on social media platforms are profiling, disinformation, echo chambers, social bots and microtargeting. Some of the risks to democracy that arise from political campaigning techniques have been highlighted in EU policies. The European Democracy Action Plan demands more transparency in political advertising and communication, specifically transparency in enforcement of the relevant rules, audits, access to non-personal data, restricting microtargeting and psychological profiling in political communication. However, the current framework, which includes the various rules on privacy, data protection, freedom of expression and free elections, is not fit for purpose. This is why this chapter argued for more precise rules on, inter alia, impact assessments including the state's duty to protect the right to free elections, the corporate responsibility to respect the right to free elections, and the right holder's access to remedies for violations of the right to free elections.

Chapter 6 suggested that the breach of the freedom of thought posed by cognitive hacking, the abstractness of the *forum internum* and the resultant intersectional dimension of the manipulation that may occur, pose an unprecedented assault on our fundamental rights. As the AI influence on independent choice increases, the human rights protections that are intended to defend free cognitive functioning must adapt to build

a fortress around individuals and their sense of self. Most laws target the protection of personal autonomy in the public sphere, while ignoring the restoration of rights in the private sphere. Existing literature indicates an abundance of macro-discourse against technological solutionism specifically in the areas of criminal profiling, e-courts and judicial due process, yet the micro-discourse of autonomous cognitive reasoning that occurs in the human mind is overlooked. This is why the chapter concluded that an update to the human rights framework, the AI Act and other legal instruments should be considered.

Part III discussed the revision of doctrines that are typically associated with private sector players, such as consent, consumer law and competition law.

Chapter 7 zoomed in on the GDPR and showed that there is a clear disconnect between the letter of the law, which positions consent as a tool for empowerment, and the reality of market actors' data processing practices. Two extreme positions on consent, hailing it either as the most important protection mechanism of data protection law or as a fundamentally misguided, flawed and unachievable concept may not be helpful to carve out a path forward. A third approach to how consent can fulfil its empowering role takes the debate outside of EU data protection law altogether and questions the business model of very large platforms, but also of much of the internet, which seems geared towards collecting as much personal data as possible for its economic exploitation. This approach would see consent continue to struggle to fulfil its empowering role so long as the business model of market actors on the internet is not fundamentally challenged. Although there are potential solutions, the chapter showed there are no easy ones.

Chapter 8 took an in-depth look at the Data Act, which aims to remove obstacles for the growth of the European data economy. The Act creates a baseline horizontal framework for compulsory data sharing. The chapter finds that it is possible to set conditions for data sharing that have a more general scope of application, but that there are several uncertainties in how they will be implemented. The success of the Data Act's horizontal framework for compulsory data sharing therefore largely depends on how its provisions are interpreted and applied. In addition, a potential pitfall is that the framework only focusses on B2B data sharing, while the horizontal framework does not apply to any pre-existing (sectoral) legislation imposing compulsory B2B data sharing. The Data Act also leaves a lot of discretion to Member States to designate competent authorities and to set up enforcement mechanisms. This might result in national differences and incompatibilities between the legal regimes of Member States.

Chapter 9 reviewed two data-driven mergers in the healthcare sector, where authorities have not explicitly engaged with non-economic data-related harms, nor with the relationship between competition and healthcare objectives. Recently, however, the calls to include public interest considerations when examining the impact of concentrations on the internal market have become louder. The chapter concluded that the focus of the

European Commission's analysis in the merger cases is somewhat limited. While there is scope for the Commission to be more proactive in imposing merger remedies that also protect non-economic interests, there are also limits to the competences of the Commission under the EU Merger Regulation. Consequently, there is a need for data protection and healthcare authorities to become involved as well. The existing framework already offers opportunities for doing so, but these have not been explored in practice yet.

Part IV considered the potential re-evaluation of doctrines that find their origin in public law, such as fairness, non-discrimination and justice.

Chapter 10 discussed the proposed AI Act and argued that it is important to examine the working of the law not just from its potential to create governance, but also for its ability to sustain and determine the knowledge around a particular issue. It explained that the AI Act should not just be understood as a piece of regulation meant to establish compliance and product liability regimes, but rather as an instrument that has the capacity to determine ways of knowing and understanding. While the chapter does not offer a template for what epistemically-just legislation should look like, it offered a contribution that demonstrates the criticality of capturing pluralistic ways of knowing in lawmaking. The chapter recommended building more reflexive lawmaking in the domain of AI, which enables greater diversity of people, experiences and expertise in shaping governance so it is representative of the plurality in our societies.

Chapter 11 discussed the contours of anti-discrimination law and showed that it is ill-suited to the context of AI decision-making on several points. This is why the chapter suggested looking at the conceptualisation of freedom as non-domination for a helpful perspective on discrimination. Although there are no clear answers, this approach could potentially account for problems revolving around intersectional forms of discrimination. It is possible to reduce intersectional discrimination through the design of AI models. In addition, non-domination may be more effective in achieving one of the fundamental goals of non-discrimination law, which is to prevent arbitrary decision-making. This approach could provide a way to reframe the problem of AI and discrimination as requiring new legal doctrine. Instead of using legal reasoning devised for analogue problems to mitigate technological and infrastructural domination, what would it mean to begin from the problems of technology? This would require new legal reasoning that addresses problems of domination, as they are created by both corporate and public technological architectures; problems of making claims in relation to our identities as groups, as individuals that experience intersecting and interacting forms of discrimination, or as subjects of pervasive technological surveillance and sorting. It would require rethinking enforcement of existing law and regulation, but also reframing those under development.

Finally, Chapter 12 assessed the ideal of legal automation, in particular in relation to judicial decision-making. Suppose that legal automation could deliver on all its expectations, would it be desirable to implement it? No, legal pragmatists like Lon Fuller would argue. The essence of law cannot be automated or, put differently, if law were to be automated, its essence, namely reciprocity, interactionality and teleology would be lost. Legal rules should be interpreted according to their meaning, not according to the text, and what purposive facts mean always requires interpretation. Even if it is clear that a matter falls within the scope of a law, its application may still be counter to its intention. To keep the essence of law, AI would need to assess whether this is a matter that should be dealt with in extra-legal ways. This requires an emotional understanding of the involved parties. Since there are often many correct legal approaches to a case, the legitimacy of a judgment depends on a party to a case having the feeling of being seen, heard and involved in the process leading up to the ultimate verdict.

## 3. A research agenda for the future

This book has yielded a high number of questions, dilemmas and complexities. Although it convincingly demonstrates by discussing data protection law, competition law, contract law, anti-discrimination law, consumer law and human rights law, that a regulatory shift may be necessary, few concrete proposals or answers are provided as of yet. This is unsurprising because the research programme at TILT has only been running for three years, but with more to follow. The reason for publishing the preliminary results is to open up this quest to everyone in the community, allowing other research groups to benefit from these results and build on them, as well as incentivising international collaborations. The main questions this book has yielded may be summarised as follows:

1. Sociotechnical change challenges traditional conceptions of law. As age-old divisions between the public and the private start to blur, should the laws governing them follow suit? To what extent should we, for example, guard privacy in public spaces (Chapter 2)? Or to what extent can private actors be enlisted in the public fight against money laundering (Chapter 3)?
2. Over the last decade, policymakers have started regulating the digital economy in earnest. However, have the concepts used as foundations for those regulations stood the test of time? Given the fast pace of socio-technical change, this is no easy feat. Take, for example, consent, which is the foundation of data protection law. Is it still effective in an economy where many platforms' business models are geared toward maximal data extraction (Chapter 7)? Similarly, does our conception of 'freedom' in non-discrimination law suffice when decision-making is increasingly automated? Are recent legislative initiatives like the Data Act capable of delivering the results expected since they remain grounded in existing paradigms (Chapter 8)?

3. Socio-technical change does not affect circumscribed areas of life and law. Rather, it is pervasive. At times, it points out the gaps between different areas of law. Can existing bodies of law uphold a siloed approach in the face of such change? Or might they have to take on values and methodologies of other bodies of law? For example, do data-driven mergers in healthcare call for a *rapprochement* between competition law, data protection and the values underlying our healthcare systems (Chapter 9)?
4. Do legislators remain capable of inclusive regulation when shaping rules like the AI Act (Chapter 10)? Are courts able to address innovative uses of technology like DNA pools for criminal investigations (Chapter 4)?
5. As private actors use technology in ways that undermine democracy and shape individual choices in illegitimate ways, how can we reshape the regulatory framework to address microtargeting, profiling and cognitive hacking (Chapters 5 and 6)?
6. Should discrimination law and by extension the human rights framework as in general be reconceptualised in light of the AI revolution and, if so, how (Chapter 11)?
7. Should law and the legal order as such be reconceptualised in light of AI and, if so, how? Or, should we ban the use of AI in the legal domain as much as possible (Chapter 12)?

Classically, regulation targets human behavior. In an era of Big Data, artificial intelligence, and robotics which shape and sometimes even replace human behavior, however, the classic regulatory paradigm is challenged in fundamental ways. This research program aims to map, understand, and – to the extent possible – help shape the shift from a human-centric regulatory paradigm to a data-centric regulatory paradigm.

This project recognizes the importance of human behavior and human decision-making, but challenges the assumption that regulation should continue to primarily be focused on facilitating adequate human decision-making. Rather, the shift towards automated decision-making—with its conflation of norm-setting and norm-enforcement—may require a more fundamental rethinking of the regulatory paradigm.

To this purpose three research lines are pursued focusing on key elements of regulation shifting from a human behaviour and human relations to data relations and the behaviour of data systems:

- how rules can be formulated and enforced,
- which concepts can be used in regulation, and
- and which values are vital.

OPEN PRESS TiU

9 789403 773148