INTERNET LAW AND DIGITAL SOCIETY

An International Overview

Edited by Paulina Kowalicka

ILS
INFORMATION, LAW & SOCIETY

Milano University Press

Contents

Introduction 13

PART I International Cybersecurity Law and Computer Crimes 15

Chapter I Challenges of Information-Sharing Under EU Cybersecurity Law: Are Competition. 17

_Hlk172573259 29

_Hlk125636801 29

Chapter II Corporate Sustainability and Digitalization of Human Resources Process 29

Chapter III Metaverse and “Meta” Crimes, Are We Facing New Threats for People Rights? 43

_Hlk124952734 45

Chapter IV Harmful Contents Online and Platforms Criminal Responsibility 49

PART II Privacy, Data Protection and Data Governance 57

Chapter V The Lies and the Fights for Privacy: Protecting Privacy and Human Dignity in the. 59

Chapter VI On the Relationship Between Competition Law and Privacy: Can we Achieve Ne. 65

Chapter VII Assessing Risks Involved in the Use of AI Systems: Current and Future Approac. 73

_Hlk125968913 86

Chapter VIII Privacy in the Digital Age: a Look at the Transformation of the Concept 89

Chapter IX The Regulation of Data Brokers in Europe: How to Address an International Data. 101

_Int_JovJxX02 105

_Int_viAJKVLK 106

_Int_29vevuQi 106

_Int_S7GIiRFL 106

_Int_foz8NXUp 106

_Int_bL6WtZxp 106

_Int_qIuZh5rB 106

_Int_FbhmTXva 106

_Int_PbepPKjl 108

Contents (continued)

Chapter X NFT: Privacy and Author Protection 111

_Int_bDRzUEGl 114

Chapter XI Digital Inheritance in Accordance with the Right to Data Protection in the Brazilia. 115

PART III Big Data, Platforms Regulations and Open Data 125

Chapter XII Legal and Ethical Challenges in the Use of Web 2.0 Open Data 127

_4kvar0bmetoi 137

_szvzmlyjkk2 137

_ofvc144m75u6 137

Chapter XIII How Smart Cities Leverage the Power of Data and Sensors to Bridge Digital Ga. 137

_nd12r1a5qn84 146

_dbatec6k8ujf 147

_nkq2cnusj6hv 147

Chapter XIV Communities' Governance in WEB3: the Role of DAOs 149

Chapter XV Revolution of Contract Through Legal Technologies: Current Trends in Contract. 161

_heading=h.d1oef056jgr 164

_heading=h.ezkgm2r569do 164

_heading=h.2dx8kbm7r7qg 164

_heading=h.j2ychfbd8vlx 164

_heading=h.9w8i7dlo6kwy 164

_heading=h.r7b974ke9kbc 164

_heading=h.fjd27j9hjbvf 165

_heading=h.i6ivwen6irh7 165

_heading=h.v67m0gf4rvnc 165

_heading=h.5xy7r3fe2gmi 165

_heading=h.2us908lge8fh 165

Contents (continued)

_heading=h.y9unj1scr7ua 165

_heading=h.utr9dj3ee6uw 165

_heading=h.wvlkm7kni7r2 165

_heading=h.xsufv8yjsjxy 165

_heading=h.g0sulkxrl8w 165

_heading=h.duqx837panw1 165

_heading=h.qxk8yb6d0kuu 165

_heading=h.o7db3p4ufdlb 165

_heading=h.ulmetr8dyyav 165

_heading=h.3xe61ah16xud 165

_heading=h.pbzixd78lqwf 165

_heading=h.6ngvucchaikz 165

_heading=h.d4m1iq213x23 165

_heading=h.4stxzcyqaeku 165

Chapter XVI “Justice” on Digital Platforms: Internal Complaint-Handling Systems and Media. 167

PART IV Artificial Intelligence, Algorithms and Legal Tech 177

Chapter XVII Opening Data of Smart Cities Under the DGA: an Overview of the Challenges. 179

_Hlk125454531 183

_Hlk125653023 187

Chapter XVIII The Algorithm in Administrative Decisions: Risks and Opportunities 191

Chapter XIX The AI Act Proposal: a New Right to Technical Interpretability? 199

_Hlk125194962 209

Chapter XX Innovative Versus Recurrent Perspectives on the Liability for Autonomous and I. 209

Chapter XXI Digitalisation of Justice in the EU, Challenges and Future Prospects 219

Chapter XXII Fairness By Design: A Value-Sensitive Approach to Exploring the Fairness Pri. 231

Contents (continued)

Chapter XXIII About the Need for Regulation Central Bank Digital Currency: Potential Mone. 241
Chapter XXIV A.I., Facial Recognition and Privacy Risk 251

INTERNET LAW AND DIGITAL SOCIETY AN INTERNATIONAL OVERVIEW

Edited By
Paulina Kowalicka

Milano University Press

Internet Law and Digital Society: An International Overview / Edited by Paulina Kowalicka. Milano: Milano University Press, 2025. (Information, Law & Society; 2)

ISBN 979-12-5510-207-6 (print)
ISBN 979-12-5510-210-6 (PDF)
ISBN 979-12-5510-212-0 (EPUB)
DOI 10.54103/infolawsoc.207

Quando non diversamente indicato, le pubblicazioni della collana Information, Law & Society sono soggette a un processo di revisione esterno, vengono valutate e approvate dal Comitato editoriale e devono essere conformi alla politica di revisione tra pari e alle indicazioni dell'Agenzia Nazionale di Valutazione del Sistema Universitario e della Ricerca.

Le edizioni digitali dell'opera sono rilasciate con licenza Creative Commons Attribution 4.0 - CC-BY-SA, il cui testo integrale è disponibile all'URL: https://creativecommons.org/licenses/by-sa/4.0

Le edizioni digitali online sono pubblicate in Open Access su: https://libri.unimi.it/index.php/milanoup.

©The Author(s), 2025
©Milano University Press per la presente edizione

Pubblicato da:
Milano University Press

Via Festa del Perdono 7 – 20122 Milano

Sito web: https://milanoup.unimi.it

e-mail: redazione.milanoup@unimi.it

L'edizione cartacea del volume può essere ordinata in tutte le librerie fisiche e online ed è distribuita da Ledizioni (www.ledizioni.it)

Table of contents

Introduction 9

PART I
INTERNATIONAL CYBERSECURITY LAW AND COMPUTER CRIMES

Chapter I
Challenges of Information-Sharing Under EU Cybersecurity Law: Are Competition Law and Data Protection Law: Hurdles or Enablers for Information-Sharing? 13
by Eyup Kun

Chapter II
Corporate Sustainability and Digitalization of Human Resources Process 25
by Claudia Ogriseg

Chapter III
Metaverse and "Meta" Crimes, Are We Facing New Threats for People Rights? 39
by Emilio Sacchi

Chapter IV
Harmful Contents Online and Platforms Criminal Responsibility 45
by Beatrice Panattoni

PART II
PRIVACY, DATA PROTECTION AND DATA GOVERNANCE

Chapter V
The Lies and the Fights for Privacy: Protecting Privacy and Human Dignity in the Digital Age 55
by Elena Pagani

Chapter VI
On the Relationship Between Competition Law and Privacy:
Can we Achieve Nexus Between Competition Law and Privacy? 61
by Arletta Gorecka

Chapter VII
Assessing Risks Involved in the Use of AI Systems: Current
and Future Approaches 69
by Pietro Boccaccini and Taís Fernanda Blauth

Chapter VIII
Privacy in the Digital Age: a Look at the Transformation of the Concept 85
by Emanuele Brambilla

Chapter IX
The Regulation of Data Brokers in Europe: How to Address
an International Data Governance Challenge 97
by Isabela Maria Rosal

Chapter X
NFT: Privacy and Author Protection 107
by Marco Alagna

Chapter XI
Digital Inheritance in Accordance with the Right to Data Protection in the
Brazilian Legal System 111
by Guilherme Vargas Puchta and Zilda Mara Consalter

PART III
BIG DATA, PLATFORMS REGULATIONS AND OPEN DATA

Chapter XII
Legal and Ethical Challenges in the Use of Web 2.0 Open Data 123
by Jonida Milaj

Chapter XIII
How Smart Cities Leverage the Power of Data and Sensors to
Bridge Digital Gaps and Foster Prosperity 133
by Beatrice Bonami

Chapter XIV
Communities' Governance in WEB3: the Role of DAOs 145
by María del Sagrario Navarro Lérida

Chapter XV
Revolution of Contract Through Legal Technologies:
Current Trends in Contract Automation 157
by Silvia Martinelli and Carlo Rossi Chauvenet

Chapter XVI
"Justice" on Digital Platforms: Internal Complaint-Handling
Systems and Mediation in P2B Relationships. A Call for Reform 163
by Ludovica Sposini

PART IV
ARTIFICIAL INTELLIGENCE, ALGORITHMS AND LEGAL TECH

Chapter XVII
Opening Data of Smart Cities Under the DGA: an Overview of the
Challenges Brought About by Data Sharing 175
by Alessandra Calvi

Chapter XVIII
The Algorithm in Administrative Decisions: Risks and Opportunities 187
by Susanna Viggiani

Chapter XIX
The AI Act Proposal: a New Right to Technical Interpretability? 195
by Chiara Gallese

Chapter XX
Innovative Versus Recurrent Perspectives on the Liability for
Autonomous and Incorporated Artificial Intelligence 205
by Juanita Goicovici

Chapter XXI
Digitalisation of Justice in the EU, Challenges and Future Prospects 215
by Anastasia Nefeli Vidaki

Chapter XXII
Fairness By Design: A Value-Sensitive Approach to Exploring the Fairness Principle in the GDPR in the Context of Children's Interaction With AI Systems 227

by Ayça Atabey

Chapter XXIII
About the Need for Regulation Central Bank Digital Currency: Potential Monetary Legal Basis and Challenges 237

by Marko Dimitrijević

Chapter XXIV
A.I., Facial Recognition and Privacy Risk 247

by Nicolò Bottura

Introduction

The Information Society Law Center (ISLC), hosted by the Department of Legal Sciences "Cesare Beccaria" of the University of Milan, is a multidisciplinary research Center, founded by Professor Giovanni Ziccardi in 2017 inside the "Cesare Beccaria" Department of Legal Sciences, devoted to the study of Legal Informatics, of Cyberspace Law and of the so called "Digital Transformation Law", alongside the legal, technological, political and social aspects of the modern Information Society.

The Research Center connects more than one hundred scholars and professionals from all over the world.

The main goals of the Center are to promote an interdisciplinary dialogue between law and new technologies, with a particular focus on the current and future changes that will deeply affect our society, and to provide valuable contributions to the development of regulatory policies capable of addressing complex issues, such as artificial intelligence, cybersecurity and cybercrime, data protection, responsible use of digital platforms and users' digital rights.

Research, education and interdisciplinary collaboration are fundamental aspects to tackle the challenges posed by new technologies, propose innovative solutions, and ensure sustainable and responsible technological development.

The future of digital policies requires forward-looking vision and bold political leadership, both at the national and international levels. It is essential to strike a balance between technological progress and the protection of fundamental rights, contributing to the creation of a future in which law serves as a key tool to guide innovation toward collective well-being and to aspire to an inclusive, safe, and beneficial digitization for all.

Milan, June 2024

Paulina Kowalicka

PART I
International Cybersecurity Law and Computer Crimes

Chapter I
Challenges of Information-Sharing Under EU Cybersecurity Law: Are Competition Law and Data Protection Law: Hurdles or Enablers for Information-Sharing?

by Eyup Kun*

INDEX: 1. Introduction. – 2. Information sharing arrangements under the NIS2 Directive. – 3. Data protection: limited certainty means limited enabler. – 4. Competition law: considering different treatment of cybersecurity service providers and non-cybersecurity service providers for enabling information-sharing. – 5. Conclusion and recommendations.

1. Introduction

Cybersecurity requires information sharing. "Cyber threats, vulnerabilities, indicators of compromise, tactics, techniques and procedures, cybersecurity alerts and configuration tools" must be shared by network and information system defenders. Sharing information about cyber threats and their spread helps prevent, detect, respond to, and mitigate incidents and improve cybersecurity. The NIS Directive (EU) 2016/1148 did not address information sharing between essential entity operators and digital service providers. NIS Directive was replaced by NIS2 Directive (Directive (EU) 2022/2555)[1] on January 16, 2023. Article 29 of the NIS2 Directive requires Member States to regulate information

* Doctoral researcher in KU Leuven Center for IT and IP Law since February 2021. He conducts his doctoral research on the intersection of cybersecurity and data protection law in the digital economy in addition to his involvement in iFLOWS and ENSURESEC project as well-founded by European Union Horizon 2020. He graduated from Istanbul University, Faculty of Law. He is a Turkish qualified lawyer since 2019. He completed his master studies at the London School of Economics and Political Science (the LSE) with the specialisation of information technology, media and communications law in 2020. During his master studies, he was involved in several projects related to the intersection between data protection and other fundamental rights. After graduating from the LSE, he worked as a trainee at the Data Protection Unit at the Council of Europe. During this assignment, he mainly worked on the guidelines on the facial recognition technologies adopted by the Council of Europe in January 2021.

1 Directive (EU) 2022/2555). Available at: https://eur-lex.europa.eu/eli/dir/2022/2555.

sharing among entities subject to the Directive and other relevant entities ("entities") for the first time at the EU level.

The underlying reason for this provision is to provide legal certainty at the EU level to encourage entities to leverage their knowledge and experience collectively to enhance their capabilities for cybersecurity. The reason can be found in recital 119 of the NIS 2 Directive[2]. Recital 119 states that entities should regularly share threat and vulnerability intelligence to detect and prevent sophisticated cyber threats. Sharing cyber threat information helps organisations prevent and recover faster. Two main legal aspects might cause uncertainty for these entities: competition and data protection law. From a competition law perspective, information sharing is based on information-sharing arrangements ("arrangements") and can be concluded among competitors. Information exchanges among competitors can be deemed illegal if it is incompatible with article 101 of the Treaty on the Functioning of the European Union (TFEU), for example, these exchanges facilitate collusive outcomes among competitors. From a data protection perspective, information-sharing practices may necessitate the processing of personal data as well as the sharing of personal data among various entities. The NIS2 Directive acknowledge the possibility of data sharing in recital 121. However, it does not provide further guidance on how these two main legal challenges can be resolved in the context of cybersecurity information-sharing practices.

The Author argues that the NIS2 Directive provides limited legal certainty for entities that engage in information-sharing. Member States should provide further certainty following the guidance given by the European Network and Information Security Agency (ENISA). For proving my argument, first, the Author will analyze the information-sharing framework under the NIS2 Directive considering its recitals and legislative discussions during its adoption. Second, the Author underlines the main data protection and competition law challenges by examining uncertainty on the potential legal basis of data processing activities in data protection law and the legality of information exchange agreements between competitors for information-sharing activities. Third, the Author will give suggestions to facilitate information sharing in cybersecurity to comply with article 101 of the TFEU and the General Data Protection Regulation (2016/679 GDPR).

2 Recital 67 and 68 of the proposed NIS2 Directive has very similar wording. See Proposal for a Directive of the European Parliament and of the Council on measures for a high common level of cybersecurity across the Union, repealing Directive (EU) 2016/1148 available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM:2020:823:FIN.

2. Information-sharing arrangements in the NIS2 Directive

Information-sharing arrangements can include information related to cybersecurity. Article 29(1) gives examples of relevant cybersecurity information: cyber threats, near misses, vulnerabilities, techniques and procedures, indicators of compromise, adversarial tactics, threat-actor-specific information, cybersecurity alerts, and recommendations about how to set up cybersecurity tools to detect cyberattacks can be shared between entities. The sharing of information could require the processing of certain types of personal data, such as IP addresses, uniform resource locators (URLs), domain names, email addresses, and, when they reveal personal data, time stamps as it is stipulated in recital 121. The need for the processing of personal data in the context of information-sharing brings uncertainty about the legality of the processing information-sharing context. To what extent the NIS2 Directive brings legal certainty regarding the processing of personal data in the context of information-sharing will be discussed in Section 3.

The NIS2 Directive promotes formal cybersecurity arrangements for information-sharing. Formal arrangements are justified for trust-building. Three factors affect information-sharing trust. First, sensitive information like a cybersecurity incident to a specific entity may be shared with other entities, which could harm its reputation[3]. Second, sharing cybersecurity information requires trust. Otherwise, they wouldn't share information to comply with legal requirements, including personal data laws. The third issue is whether information sharing is two-way. Thus, these entities sharing information should benefit all parties. Article 26(2) of the proposed NIS2 Directive calls these entities "trusted communities", supporting the role of trust in arrangements. The NIS2 Directive removed it, but it does not change the motivation for formal arrangements. Thus, the NIS2 Directive's cybersecurity contractual framework seems justified.

Entities falling within the scope of the NIS2 Directive and other entities, where relevant, can be a party to the arrangements according to article 29 of the NIS Directive. Regarding the entities falling within the scope of the NIS2 Directive, article 2 of the NIS2 Directive determines the scope of the NIS 2 Directive[4]. Regarding "the scope of other entities", the scope of these entities is all types of entities that are relevant for information sharing. In particular, relevant entities are cybersecurity service providers and entities that focus on

3 Agrafiotis, I., et al. (2018) 'A Taxonomy of Cyber-Harms: Defining the Impacts of Cyber-Attacks and Understanding How They Propagate', *Journal of Cybersecurity*, vol. 4(1). Available at: https://academic.oup.com/cybersecurity/article-abstract/4/1/tyy006/5133288/.

4 The scope of entities that are responsible under the NIS2 Directive is substantially extended. See article 2 and 3, and different sectors (such as energy, transport, health, financial market infrastructure).

cybersecurity research. Initially, the proposed version of the NIS2 Directive does not stipulate these entities as part of the arrangements[5]. The proposed NIS2 refers to the possibility of extension of the information arrangements to other relevant entities in its recital 68 without providing any further examples. In the final version of the NIS2 Directive, it is extended and further clarified by two additions to article 29 and an insertion to recital 120. Regarding the additions to the provision, first, "where relevant, other entities not falling within the scope of Directive" is included in article 29(1). Second, "where relevant, their suppliers or service providers" is added to article 29(2). In addition to Recital 120, "such as those providing cybersecurity services and research" is inserted. These clarifications help the involvement of different stakeholders in cybersecurity information sharing to be part of these arrangements. Because the entities subject to NIS2 Directive can outsource their cybersecurity services to other relevant entities, it is reasonable to insert these entities directly into the arrangements to provide an opportunity to share their practical experience with other entities. While the entities are party to these arrangements voluntarily, the NIS 2 Directive requires important and essential entities to notify their relevant competent authorities of their participation as well as withdrawal from these arrangements according to article 29(4) of the NIS 2 Directive.

Sharing information with competitors might help ensure cybersecurity since they are likely to address similar types of cybersecurity attacks or challenges. The information exchanges between competitors should be assessed in the competition law since it might reveal sensitive information regarding their market strategies. Thus, competition law may be assessed to ensure that cybersecurity information-sharing among competitors respects competition law in particular article 101 of the TFEU. Section 4 addresses this issue.

3. Data protection: limited certainty means limited enabler

In the case of information-sharing, one of the most pressing issues is to ensure that the information-sharing is compatible with the GDPR. The compatibility means the information-sharing complies with the GDPR requirements. While there are a variety of requirements to comply with the GDPR such as data subject rights and its limits, international data transfers[6], and transparency of the processing of personal data, this paper discusses the legal basis for the

5 See article 26 of the proposed NIS2 Directive.

6 See Cormack, A. (2021) 'NISD2: A Common Framework for Information Sharing among Network Defenders', *SCRIPTed: A Journal of Law, Technology and Society*, vol 18(1). Available at:https://script-ed.org/article/nisd2-a-common-framework-for-information-sharing-among-network-defenders/.

processing of data in particular sharing the personal data with other entities. The NIS2 Directive stipulates legitimate interest as a legal basis for data-sharing in recital 121. The Author contends that the clarification provides only limited legal certainty for the legal basis for information-sharing at the EU level, which may impede information-sharing among various entities. It means failing to meet the goal of the information-sharing framework. To support my claim, the Author discusses the legal bases under article 6(1) of the GDPR that are relevant in the context of information sharing. Following that, the Author will demonstrate the NIS 2 Directive's limits on certainty.

For the arrangements, the most relevant legal grounds might be consent (article 6(1)(a)), the legal obligation (article 6(1)(c)), the necessity for the performance of a task carried out in the public interest (article 6(1)(e)) as well as a legitimate interest (article 6(1)(f)). Regarding consent as a legal basis, consent may not be an appropriate legal basis for cybersecurity information sharing for two reasons. First, data subjects can revoke consent at any time under article 7(3) of the GDPR[7]. Personal data processing for information-sharing should not depend on consent withdrawal. This dependency may make information-sharing uncertain. Second, there is a high probability that this won't cover the kind of threat intelligence that needs to be shared. For instance, it is more likely that the user of the IP address is an independent bad actor rather than a person with whom the company already has an established relationship[8].

Regarding the legal obligation basis in the context of cybersecurity, it can be argued that the NIS 2 Directive establishes a legal basis for the processing of personal data for cybersecurity purposes. Article 2(14) of the final version of the NIS 2 Directive states that entities must process personal data to the extent necessary for this Directive and per GDPR, specifically article 6. Similarly, recital 121 of the NIS 2 Directive clarifies that essential and important entities may process personal data to the extent necessary and proportionate to secure network and information systems following article 6(1), point (c) and article 6(3) of Regulation (EU) 2016/679. However, as previously stated, cybersecurity information-sharing is not considered an obligation for entities. These entities can participate in information-sharing activities voluntarily. As a result, entities

7 For a similar argument, see Albakri, A., Boiten, E., Lemos, R. (2019) 'Sharing Cyber Threat Intelligence Under the General Data Protection Regulation', *Lecture Notes in Computer Science. Privacy Technologies and Policy,* vol. 11498. Available at: https://link.springer.com/chapter/10.1007/978-3-030-21752-5_3; Sullivan, C., Burger, E. (2017) "In the Public Interest': The Privacy Implications of International Business-to-Business Sharing of Cyber-Threat Intelligence', *Computer Law & Security Review,* vol. 33(1). Available at: https://doi.org/10.1016/j.clsr.2016.11.015.

8 Sullivan, C., Burger, E. (2017) "In the Public Interest': The Privacy Implications of International Business-to-Business Sharing of Cyber-Threat Intelligence', *Computer Law & Security Review,* vol. 33(1). Available at: https://doi.org/10.1016/j.clsr.2016.11.015.

cannot rely on the NIS 2 Directive's legal obligation to process personal data for information-sharing purposes.

Regarding reliance on the public interest under article 6(1) e, according to this provision, "personal data shall be processed if the processing is necessary for the performance of a task carried out in public interest or the exercise of official authority vested in the controller". By conceptualising information-sharing in the public interest, Sullivan and Burger argued that data controllers can rely on that legal basis when they deploy automated sharing of IP addresses[9]. Considering the importance of information sharing for collectively defending the cyber-sphere, it is true that information sharing indeed falls within the scope of public interest.

However, the Author does not agree with Sullivan and Burger on the reliance on article 6(1)e on the following reasoning. It is unclear whether the words "vested in the controller" refer to "exercise of official authority" or "a task" in the English version of article 6(1)(e)[10]. As Kotchy argues that the German version, where commas are used to structure the sentence, clarifies the meaning[11]. This structure would be translated into English as follows: "Processing is required for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller". However, Sullivan seems to interpret the first element "processing is necessary for the performance of a task carried out in public interest" without considering the second element "vested in the controller"[12]. Therefore, as long as entities as data controllers are not vested in a task of information-sharing with other entities, they cannot rely on the article 6(1)(e). As another supporting point, in *Meta* case, the Court of Justice of the European Union places particular emphasis on the "entrusted with a task" aspect when Meta asserts that it conducts research for the purpose of "social good, and to promote safety, integrity, and security" and can invoke article 6(1)(e) of the GDPR. According to the CJEU, it is unlikely that a private operator would be given the responsibility for such a task, considering the nature of the activity and its primarily economic and commercial nature[13]. Entities under the NIS 2 Directive cannot be considered as "entrusted with a task" in the in the information-sharing since the information-sharing is only a possibility for them. Therefore, they cannot rely on the legal basis under article 6(1)(e) of the GDPR.

---

9 Ibid.

10 Kotschy, W. (2020) 'Article 6 Lawfulness of Processing', *Oxford University Press*. Available at: https://oxford.universitypressscholarship.com/10.1093/oso/9780198826491.001.0001/isbn-9780198826491-book-part-35.

11 Ibid.

12 Ibid.

13 The Court of Justice of the European Union, Meta Platforms Inc. v. Bundeskartellamt (C-252/21), ECLI:EU:C:2023:537, paragraph 133.

Regarding the legitimate interest, personal data processing is lawful only if it is required for the controller's or a third party's legitimate interests unless such interests are overridden by the data subject's interests, fundamental rights and freedoms according to article 6(1)(f). As a result, that provision establishes three cumulative conditions for the lawfulness of personal data processing: the pursuit of a legitimate interest by the data controller or a third party, the need to process personal data for the legitimate interests pursued; and the interests, freedoms, and fundamental rights of the person concerned by data protection do not take precedence. Cybersecurity is one of the legitimate interests explicitly recognised under recital 49 of the GDPR for providers of security technologies and services. In particular, data controllers inevitably process personal data for the prevention, detection and investigation of security incident security[14]. Cormack discovered a strong alignment and proposed pertinent factors for the legitimate interests balancing test to guarantee that the interests of users were protected when they shared information[15]. Similarly, Bakri and others argued for the appropriateness of legitimate interest by proposing data protection by design approach to balance the interests of data subjects and the interest to share information[16]. While the GDPR does not make specific reference to information-sharing in the context of cybersecurity, the NIS2 Directive adds clarification by making specific reference to information-sharing as a legitimate interest for personal data processing in recital 121[17]. Recital 121 states that when personal data processing is required for voluntary information-sharing among entities, this processing is permissible. By stating that information-sharing can be considered necessary for legitimate interest, the NIS2 Directive incorporates information-sharing into the framework of legitimate interest and supplements recital 49 of the GDPR[18].

---

14 CIPL Publishes White Paper on How the Legitimate Interest Ground for Processing Enables Responsible Data Use and Innovation (2021). *Privacy & Information Security Law Blog*. Available at: https://www.huntonprivacyblog.com/2021/07/21/cipl-publishes-white-paper-on-how-the-legitimate-interest-ground-for-processing-enables-responsible-data-use-and-innovation/.

15 Cormack, A. (2016) 'Incident Response: Protecting Individual Rights Under the General Data Protection Regulation', *SCRIPTed: A Journal of Law, Technology & Society*, vol. 13(3). Available at: https://script-ed.org/article/incident-response-protecting-individual-rights-under-the-general-data-protection-regulation/.

16 See Albakri, A., Boiten, E., Lemos, R. (2019) 'Sharing Cyber Threat Intelligence Under the General Data Protection Regulation', *Lecture Notes in Computer Science. Privacy Technologies and Policy,* vol. 11498. See for another approach, Von Maltzan S. (2019) 'No Contradiction Between Cyber-Security and Data Protection? Designing a Data Protecton Compliant Incident Response System', *European Journal of Law and Technology*, vol. 10(1). Available at: https://ejlt.org/index.php/ejlt/article/view/665.

17 Recital 69 of the proposed NIS2 Directive has similar reference to legitimate interest.

18 Cormack, A. (2021) 'NISD2: A Common Framework for Information Sharing among Network Defenders', *SCRIPTed: A Journal of Law*, *Technology and Society*, vol 18(1). Available at:

The legal certainty provided by the NIS2 Directive is limited because it only refers to legitimate interest as a legal basis for the processing of personal data in the context of information-sharing and clarifies the nature of legitimate interest and necessity for the processing in the three-step legitimate interest test. It does not provide any further guidance on the proportionality aspects of personal data processing in the context of arrangements or interests of data subject rights.

4. Competition law: considering different treatment of cybersecurity service providers and on-cybersecurity service providers for enabling information-sharing

Because arrangements can be made between competitors, competition law concerns about information sharing between entities might be uncertain. In the context of cybersecurity information sharing within article 29 of the NIS2 Directive, the Author argues that compliance with article 101 of the TFEU should be analyzed in two different contexts: information-sharing between entities that are not providing cybersecurity services and information-sharing between entities that provide cybersecurity services. Considering these two different entities and markets, the Author maintains that the guidance should be differentiated for these two different entities. To prove my argument, first, the Author gives a brief overview of the article 101 of the TFEU and the information exchange arrangements in particular. Second, the Author analyzes these two different contexts within the framework the Author described.

One of the goals of article 101 of the TFEU is to ensure that undertakings do not use horizontal cooperation agreements to prevent, restrict, or distort market competition to the detriment of consumers[19]. Article 101 does not apply to horizontal cooperation unless there is some form of coordination between competitors, such as an agreement between undertakings, a decision by an association of undertakings, or a concerted practice. The existence of an agreement, a concerted practice, or a decision by a group of businesses does not imply that there is a restriction on competition under article 101(1). Article 101 evaluation consists of two steps. The first step, according to article 101(1), is to determine whether an agreement between undertakings that has the potential to affect trade between Member States has an anti-competitive object[20] or

https://script-ed.org/article/nisd2-a-common-framework-for-information-sharing-among-network-defenders/.

19 Article 101 of the TFEU.

20 Guidelines on the applicability of Article 101 of the Treaty on the Functioning of the European Union to horizontal co-operation agreements, available at: https://competition-policy.ec.europa.eu/document/fd641c1e-7415-4e60-ac21-7ab3e72045d2_en, (Horizontal Guidelines).

actual or potential restrictive effects on competition[21]. Article 101(3) says that the second step is only important if an agreement is found to be "restrictive of competition" in the sense of article 101(1), is to find out what benefits the agreement has for competition and decide if these benefits outweigh the negative effects on competition[22]. Horizontal cooperation can be a means to share risk, save costs and enhance product quality and variety. Horizontal cooperation agreements may however also limit competition in several ways. The agreement may for instance lead to a loss of competition on the relevant market, risk of collusion between the parties or foreclosure[23].

The information-sharing arrangements can fall within the scope of article 101 of the TFEU. Regulatory initiatives may also result in information exchange. Even if undertakings are encouraged or required to share certain information and data to meet Union or government requirements, article 101(1) remains in effect. Information exchange is a common feature of many competitive markets and may generate various types of efficiency gains[24]. Data sharing has gained importance through the use of big data analytics and machine learning techniques[25]. Undertakings should avoid exchanges of information that have the object or effect to give rise to conditions of competition which do not correspond to the normal conditions of the relevant market. Information exchange can enable undertakings to achieve a collusive outcome on markets where they would otherwise not have been able to do so. It can also lead to anti-competitive foreclosure on the same market where the exchange takes place or on a related market[26]. This type of foreclosure is possible if the information concerned is of strategic importance and covers a significant part of the relevant market.

Article 101(1) applies if an exchange of commercially sensitive information is likely to influence the commercial strategy of competitors. This is the case if information reduces uncertainty regarding one or several competitors' future

---

Article 101(1) prohibits both actual and potential anti-competitive effects; see for example judgment of 28 May 1998, *John Deere*, C-7/95 P, EU:C:1998:256, paragraph 77.

21 Object v. effect restriction citation.

22 Horizontal Guidelines. See judgment of 6 October 2009, *GlaxoSmithKline*, C-501/06 P, C-513/06 P, C-515/06 P and C-519/06 P, EU:C:2009:610, paragraph 95.

23 Horizontal Guidelines.

24 Gonzàlez, A.O. (2012) 'Object Analysis in Information Exchange among Competitors'. Available at: https://www.semanticscholar.org/paper/Object-analysis-in-information-exchange-amongGonz%C3%83%C2%A1lez/4bb30af6a40e407cfb12e2976ccc815686139b3a; Ferretti, F., (2014) 'Information Exchanges Under EU Competition Law', *Springer International Publishing*. Available at: http://link.springer.com/10.1007/978-3-319-08906-5_3.

25 See for general analysis of framework that applies to data sharing, Graef, I., Tombal, T., Streel, A., (2019) 'Limits and Enablers of Data Sharing. An Analytical Framework for EU Competition, Data Protection and Consumer Law', *Electronic Journal, TILEC Discussion,* Paper No. DP 2019-024. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3494212#.

26 Horizontal Guidelines.

or recent actions in the market and regardless of whether the undertakings involved in the exchange obtain some benefit from their cooperation.

In the context of cybersecurity information sharing within Article 29 of the NIS2 Directive, compliance with article 101 of the TFEU should be analyzed in two different contexts: information-sharing between entities that are not providing cybersecurity services and information-sharing between entities that provide cybersecurity services.

Concerning the former, competition law concerns about information exchange do not pose a significant risk because cybersecurity information is not considered commercially sensitive[27]. The following example demonstrates this in practice. For example, online marketplaces as digital providers (Annex II "Other Critical Sectors") are covered by the NIS2 Directive. Their business model is to provide a platform which facilitates the transaction between businesses and consumers. These entities may be included in information-sharing agreements. When they share relevant cybersecurity information among themselves, such as cyber threats, near misses, vulnerabilities, techniques and procedures, and indicators of compromise, this information cannot be considered commercially sensitive for their services because it is not directly related to their commercial activities or strategies. As a result, it is less likely that this type of information has resulted in either object or effect-based competition restriction.

Regarding the latter, the Author argues that further guidance might be needed for information exchanges between competitors in the sector of cybersecurity service since it might pose a risk to compliance with article 101 of the TFEU. Despite the consideration of information sharing among the entities as public interest, threat intelligence, which is a form of information-sharing regarding cybersecurity is commercially exploited. Threat intelligence products and services in cybersecurity market inform cybersecurity threats and other issues to different entities in different services. These products and service providers help curate information about threats' identities, motivations, characteristics, tactics, techniques, and procedures (TTPs). Better decision-making and security technology capabilities reduce risk and compromise[28]. The NIS2 Directive anticipates cybersecurity service providers participating in arrangements. Within the NIS2 Directive, threat intelligence providers may be considered relevant "other entities". When these entities use arrangements to share specific TTPs or know-how with their competitors, the information they provide may be commercially sensitive within this specific market. The information is likely to

27 FTC, DOJ Issue Antitrust Policy Statement on Sharing Cybersecurity Information (Federal Trade Commission, 10 April 2014) Available at: https://www.ftc.gov/news-events/news/press-releases/2014/04/ftc-doj-issue-antitrust-policy-statement-sharing-cybersecurity-information.

28 Janeja, V.P. (2022) 'Data Analytics for Cybersecurity', Cambridge University Press, p. 12.

be used strategically within the sector to anticipate future products[29] and might result in collusive practices within the threat intelligence sector. The likelihood of collusion is also depending on the market characteristics[30]. This information exchange should restrict competition by object or by effect considering legal and economic context to fall within article 101 of the TFEU[31]. In terms of the economic and legal context, because the arrangements contained in the NIS 2 Directive aim to ensure a collective response to cybersecurity concerns, these arrangements may not be regarded as a restriction by object or effect. However, the analysis should be done on a case-by-case basis, taking into account all aspects of the arrangements.

5. Conclusion and recommendations

The NIS2 Directive aims to facilitate information-sharing activities between entities to increase the level of cybersecurity in the EU. The underlying reason for regulating in the NIS 2 Directive is to provide legal certainty. In particular, the NIS 2 Directive refers to uncertainties related to competition law and data protection law. I argue that although the NIS2 Directive provides legal limited certainty, it must be further substantiated by the Member States in light of guidance given by the ENISA at the EU level. Otherwise, uncertainties related to data protection and competition law might hinder information-sharing among different entities, which fails to meet the objective of the information-sharing.

The NIS2 Directive places obligations on Member States and ENISA to facilitate arrangements between these entities. According to article 29(4), Member States must specify operational elements as well as the content and conditions of the arrangements. This type of facilitation by Member States is critical in encouraging information-sharing practices. It is critical because uncertainty about the legal requirements may cause entities to refrain from participating in information-sharing practices. If Member States alleviate those concerns by establishing legal standards for data protection and competition, as discussed in Sections 3 and 4, they will make it easier for entities to engage in those activities. Because the NIS2 Directive seeks minimal harmonisation under article 5, and it is a directive, Member States have the liberty to require different operational elements and conditions for information-sharing arrangements. However, because of the various requirements and conditions, entities may choose not to participate in information-sharing arrangements.

To prevent divergences at the Member States' level, ENISA shall guide at the EU level. Article 29(5) of the NIS Directive requires ENISA to provide

29 Horizontal Guidelines.
30 Ibid.
31 Ibid.

guidance and exchange good practices at the EU level to avoid the problem of disincentivizing divergences regarding the operational elements as well as the conditions of these arrangements. ENISA should guide Member States while they provide further certainty to ensure that there are no diverging approaches in data protection and competition law aspects of information-sharing arrangements.

Regarding data protection uncertainties, in particular legal basis, if Member States do not provide further guidance on the legal basis in particular concerning the balancing test in the legitimate interest in the transposition phase of NIS 2 Directive, the Directive would not meet its objective of enabling information-sharing among the different entities. The Author proposes that Member States should define which type of personal data can be shared within the arrangements by creating a non-exclusive list of data that might be legally shared within these arrangements. For instance, whether IP addresses[32] and URLs that can be considered personal data can be shared within the arrangements should be clarified. In addition, while providing operational guidance for the arrangements, Member States should guide specific safeguards such as data protection by design requirements within specific platforms where information sharing may occur. Guiding through specific safeguards will assist entities to balance the data protection rights of data subjects with the information-sharing interests.

Regarding competition law, it appears that information sharing does not pose a legal challenge for entities that do not provide cybersecurity as a service. However, information exchange within cybersecurity service providers may pose a risk of non-compliance with article 101 of the TFEU if commercially sensitive information is not properly shared. This issue may deter those entities from participating in arrangements. There is no also specific guidance given to the arrangements in cybersecurity in the Horizontal Guidelines, adopted by the EU Commission in June 2023.

As a suggestion, while guiding operational aspects of arrangements, Member States should consider competition concerns regarding information exchange when designing those arrangements. In particular, the conditions of information access for cybersecurity service providers that may be considered commercially sensitive should be appropriately designed to prevent these service providers from gaining additional insight into specific products or services of other competitors. The second suggestion is that Member States ensure that cybersecurity service providers have non-discriminatory access to those arrangements to reduce the risk of foreclosure.

32 See for the case-law analysis of IP address, Sullivan, C., Burger, E. (2017) "In the Public Interest": The Privacy Implications of International Business-to-Business Sharing of Cyber-Threat Intelligence, *Computer Law & Security Review,* vol. 33(1). Available at: https://doi.org/10.1016/j.clsr.2016.11.015.

Chapter II
Corporate Sustainability and Digitalization of Human Resources Process

by Claudia Ogriseg*

Index: 1. Sustainable Development Goals (SDGs) and corporate sustainability in digital transformation. – 2. The European strategy for the development of corporate sustainability. – 3. Critical aspects for the development of corporate sustainability in digitized/digital companies or industry implemented by 4.0 technologies. – 4. The sustainability of algorithmic management: transparency, fairness and accountability in the EU Commission proposal for artificial intelligence and work on platform. – 5. Promoting corporate sustainability through the corporate disclosure.

1. Sustainable Development Goals (SDGs) and corporate sustainability in digital transformation

In 2020, the European Commission defined an industrial strategy to support a dual transformation to a green and digital economy and digital transformation received a strong acceleration during the pandemic. Digital transformation in companies has been a shift towards efficiency, innovation, competitiveness, sustainability and overall economic growth of the organizations. It refers to unprecedented disruptions to society, industry, organizations stimulated by advances in digital technologies such as: artificial intelligence, big data analytics, cloud computing and the Internet of Things.

The development of corporate sustainability is the result of an improvement process that considers profits, environment and people[1]. Critical issues in sus-

---

* Labour Lawyer in Italy she obtained a PhD in Labor Law and Industrial Relations at the University of Bologna and deepened the issues of personal data protection and digital work as Research Fellow at the Multidisciplinary Research Center Information Society Law Center (ISLC) at the University of Milan. She is currently Head and Tutor for the Regional branch of the School for Higher Education in Labor, Trade Union and Social Security Law "Luca Boneschi" and Data Protection Officer (DPO) for public bodies and medical facilities. Her research investigates issues relevant to human resource management in companies facing paths of digitization or use of artificial intelligence systems.

1 Companies' sustainability is "meeting the needs of a firm's direct and indirect stakeholders (such as shareholders, employees, clients, communities, etc.), without compromising its ability to meet the needs of future stakeholders as well." Dyllick, T., Hockerts, K. (2022) 'Beyond

tainable development paths arise from the difficulty of measuring, monitoring and communicating the effects of corporate actions in social terms. This article aims to analyse the measurement, management and reporting on corporate sustainability in the digital transformation.

2. The European strategy for the development of corporate sustainability

The EU's focus on sustainability issues has led to the provision of increasingly important regulation in no-financial reporting[2]. Promoting a Sustainable Financial system was the first step in ensuring welfare, social inclusion, and reduced exploitation of natural resources and the environment. In the Directive no. 2014/95/EU No Financial Reporting Directive (NFRD), the main regulatory intervention on no-financial reporting, the European Union requires macro-categories of data to report on gender gap the annulment policies and their results, the respect for workers' rights and their working conditions, diversity and inclusion policies, including with regard to corporate bodies.

Given the extreme generality of the Directive's requirements and the existence of a large number of standards and frameworks, there has always been a great deal of uncertainty about the standard which best suits the needs and expectations of stakeholders. To ensure greater coverage, evaluation and transparency corporate companies even adopt more than one standard for disclosure: the Global Reporting Initiative (GRI)[3], the Sustainability Accounting Standards Board (SASB)[4], the IFRS Sustainability Disclosure Standards developed by the International Sustainability Standards Board (ISSB)[5], the Standard ISO[6] etc. Downstream of the corporate strategy, the most complex intervention in the development of sustainability concerned the selection of the best Key Performance Indicators (KPIs) for monitoring performance relative to corporate objectives and the performance levels the company wishes to achieve.

To increase the quality of corporate disclosure and guide in the correct implementation of the Directive no. 2014/95/EU No Financial Reporting Directive, the European Commission drew up guidelines on the communication

---

the Business Case for Corporate Sustainability' *Business Strategy and the Environment*, vol. 11, pp. 130-141. Available at: https://doi.org/10.1002/bse.323

2 Minutiello, V., Brunello, L. (2022) *Voluntary vs Mandatory Disclosure delle performance di sostenibilità,* in Tettamanzi, P., V. Minutiello, *ESG: bilancio di sostenibilità e integrated reporting*, Wolters Kluwer Ipsoa, vol. 105.

3 See https://www.globalreporting.org/.

4 See https://sasb.org/.

5 See https://www.ifrs.org/content/dam/ifrs/project/general-sustainability-related-disclosures/effects-analysis.pdf.

6 See https://www.iso.org/standards.html.

of non-financial information (Commission 2017/C 215/01). The Commission does not define a reporting standard: it refers to some of the main european and international standards of non-financial information including, for example, the Global Reporting Initiative (GRI) and the Standard ISO 14000. Regarding the social aspects of the report, the Commission provides some examples focusing on workers' rights, dialogue with trade unions, respect for diversity within the organisation, management arrangements and employee engagement, staff training activities, and employee and consumer health and safety.

The first Action plan for a sustainable growth for the European Sustainable Finance was defined in 2018. The EU Taxonomy was adopted with the EU Regulation no. 2020/852 and the Group of Expert – assisting the European Commission – presented a Final Report on Social taxonomy[7].

Directive no. 2022/2464/EU Corporate Sustainability Reporting Directive (CSRD) provided significant innovations: it modified the Directive no. 2013/34/EU (concerning the obligation to communicate information of non-financial nature for large companies) and expanded sustainability disclosure through the current Directive no. 2014/95/EU Non-Financial Reporting Directive (NFRD). The CSRD started a process of EU standardisation on the Sustainability Report in all fields: environmental, social, governance (ESG). It values sustainability metrics alongside environmental performance, paying particular attention to the "S" of "ESG" (employee health, human rights, corruption, anti-corruption, diversity). The standards will reflect the disclosure requirements defining from EU (i.e. green taxonomy, the european pillar on social rights and the sustainable corporate governance and due diligence directive) and will be directly issued by the European Commission in ad hoc measures, so-called "Delegated Acts".

Sustainability reports will be subject to "limited assurance", with the aim of achieving the "reasonable assurance" (which is typical of the economic-financial balance) and their revision must be carried out by an accredited «Statutory auditor». Companies will be obliged to include sustainability reports in the Management Report. To ensure greater comparability between disclosures, companies will be required to adopt a single reporting standards ESRS (European Sustainability Reporting Standard), whose development is mandated to the European Financial Reporting Advisory Group (EFRAG). The reporting standards developed by the EFRAG (ESRS), compared to the GRI and SASB indicators, are aimed specifically at European companies and aligned with the EU principles on sustainability.

Last 31 July 2023 the European Commission adopted the first set of application standards to enable companies to comply with Directive no.2022/2464/

7 The EU Regulation no.2020/852 has established the Taxonomy, that is the unified system of classification of sustainable economic activities in Europe aimed at encouraging investments with environmental and social objectives.

EU Corporate Sustainability Reporting Directive ESRS obligations[8]. The standards shall reflect t reporting requirements (contained in the EU Green Taxonomy and in the proposal of the Directive on Corporate Sustainability Due Diligence CSDD) and will be crucial to understand whether the corporate sustainability can also be effectively protected from non-financial information disclosure requirements. The quality of Integrated Reporting (IR) and Sustainability Reporting (SR) could be important in reducing information asymmetry and ensure a real social sustainability in organizations. The first set consists of 12 standards: two Cross Cutting General Scope Standards and Ten Topical Standards (Environmental, Social, Governance). The company will be required to report on working conditions, access to equal opportunities and other labour-related rights regarding the workforce (ESRS1) and supply chain (ESRS2). Information on the impact of the company's operations and value chain, including its products and services on local communities, on civil, social and economic rights, including water and sanitation, relevant to local communities should be assessed (ESRS3). Finally, information on the impact of a company's products and/or services on consumers and end-users, including access to quality, privacy and child protection information (ESRS4), should be collected and made transparent.

On 22/12/2023 the Delegated Regulation (EU) 2023/2772 of 31 July 2023, supplementing Directive 2013/34/EU of the European Parliament and of the Council with regard to sustainability reporting principles, was published in the Official Journal of the European Union[9]. The Regulation describes the social, environmental and governance information on which companies will have to report from 1 January 2024 and for all subsequent financial years. Reporting in accordance with the European Sustainability Reporting Standard (ESRS) should enable users to understand the significant impacts of the undertaking on people and the environment and the effects on the development, performance and situation of the undertaking. The purpose of these reporting criteria is to communicate the contribution to sustainable development by identifying the impacts that the company has or may have on the environment and people (including human rights impacts related to the business and the supply chain), financial risks and opportunities arising from dependence on natural, human and social resources, identified through a financial significance assessment process.

The company will be required to disclosure the information according to the principle of double relevance of impact and financial (Reg. Chapter 3). Sustainability issues will be reported to the extent that they are impact-relevant (i. e. have significant effects in the medium or long term on the environment, people, the company's own activities and/or the value chain, including through

8 See https://finance.ec.europa.eu/news/commission-adopts-european-sustainability-reporting-standards-2023-07-31_en.

9 See https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202302772.

its products and services and its business relationships) and relevant to financial profiles. (i. e. capable of generating risks or opportunities that affect or can reasonably be expected to affect the entity's financial position, profit or loss, cash flows, access to finance or cost of capital in the short, medium or long term).

In the landscape of directives and measures that the European Union is adopting on sustainability reporting (the CSRD directive, the new ESRS standards etc.), the proposed Corporate Sustainability Due Diligence Directive is also included. The proposal for a Directive on Corporate Sustainability Due Diligence (CSDD)[10] focuses on the due-diligence duties of a company and its impact on human rights including workers' rights, health, climate, environment. Companies will have to identify risks and information will be required on company policy; company processes; company activities; company findings; measures taken by the company; outcomes of these measures. Identify social objectives symmetrical to the environmental set and define a social taxonomy will allow to measure and report a quali-quantitative social impact. Guarantee a real mandatory disclosure on personal data processing profiles and access to human oversight and eventual correction of the system will enable decent work, a sustainable protection of the individual through the remedies of data protection, non-discrimination, labour law.

3. Critical aspects for the development of corporate sustainability in digitized/digital companies or industry implemented by 4.0 technologies

In the EU sustainability strategy, companies will commit to integrating ESG objectives into their strategy and will need to communicate how sustainability initiatives affect the company's performance, its results, the financial situation and the structure of the business model. Within the risk management model (ERM – Enterprise Risk Management), companies will be required to consider climate-related risks and other environmental issues, such as biodiversity loss and health and social issues, including child labor and forced labor.

In the social fields, ESRS S1, regarding the own workforce, aims to specify disclosure requirements that will enable users of the sustainability statement to understand the company's material impacts on its workforce, as well as related risks and opportunities.

10 Amendments adopted by the European Parliament on 1st June 2023 on the proposal for a Directive of the European Parliament and of the Council on Corporate Sustainability Due Diligence and amending Directive (EU) 2019/1937 (COM (2022)0071 – C9-0050/2022 – 2022/0051(COD)). Available at: https://www.europarl.europa.eu/doceo/document/TA-9-2023-06-01_EN.html#sdocta4.

The disclosure requirements should consider strategies (interest and views of stakeholders, material impacts, risk and opportunities and their interaction with strategy and business model). In order to achieve the objective, the ESRS S1-S2-S3 Standard also requires an explanation of the general approach the company takes to identify and manage any actual and potential material impacts on its workforce/value chain workers/affected communities in relation to the following social factors or issues, including human rights and information regarding impact, risks and opportunities management policies related to its workforce/value chain workers/affected communities (S1-1; S2-1; S3-1), processes for engaging with own workers'/value chain workers representatives about impacts (S1-2; S2-1; S3-2) or to remediate negative impacts and channels for its workers/value chain workers to raise concern (S1-3; S2-3; S3-3); taking action on material impacts on its workforce/value chain workers/affected communities, and approaches to mitigating material risks and pursuing material opportunities related to own workforce/value chain workers/affected communities, and effectiveness of those actions (S1-4; S2-4; S3-4).

The disclosure requirements are targets related to own workforce/value chain workers in managing relevant negative impacts, advancing positive impacts, and managing relevant risks and opportunities (S1-5; S2-5) and about its own workforce characteristics of the undertaking's employees (S1-6); characteristics of non-employee workers in the undertaking's own workforce/value chain workers (S1-7); collective bargaining coverage and social dialogue (S1-8); diversity metrics (S1-9); adequate wages(S1-10); social protection (S1-11); persons with disabilities (S1-12); training and skills development metrics (S1-13); health and safety metrics (S1-14); work-life balance metrics (S1-15); compensation metrics (pay gap and total compensation) (S1-16); incidents, complaints and severe human rights impacts (S1-17).

The aim of the Standards (S1-S2-S3) is also to enable users to understand the extent to which the company aligns or complies with international and European human rights instruments and conventions, including the EU labour law acquis.

The metrics consider working conditions (including secure and adaptable employment, working time, adequate wages, social dialogue, collective bargaining and the involvement of workers, work-life balance, a healthy, safe and well-adapted work environment) respect for human rights and fundamental freedoms included in the Charter of Fundamental Rights of the European Union (freedom of association, works council, consultation and participation rights workers), equal treatment and opportunities (gender equality and equal pay for equal work, training and skills development, employment and inclusion of people with disabilities, measures against violence and harassment in the workplace) other work-related issues, including child labour, forced labour adequate housing and privacy/data protection.

Any sustainability journey involves not only the identification of a system of indicators, but also the collection and analysis of data as well as the internal and external communication of results with the preparation of sustainability reports. Sustainability reports demonstrate the commitment of the company, generate awareness in internal (employees) and external (regulators, suppliers, employees in the value chain) stakeholders.

For companies involved in digital transformation on one hand, digitization of production processes including the use of wearable devices (smart watches, belts and gloves) can contribute to a healthier and safer workplace. On the other hand, machine/equipment monitoring, and smart worker tracking enable companies to collect massive information (big data), profile employees and artificial intelligence let company to combine data so as to make automated decisions on collaboration, task organization and labour productivity.

Digital enterprises, which aim in measuring their social sustainability, need to estimate KPIs on healthy and safe working conditions even regarding mental status, on a real and transparent protection of personal data and private information. Mapping issues related to the social sustainability in the digital transition highlights an increase in the processing of personal data, which is likely to become more complex, potentially pervasive and "harmful" to human rights. Digital transformation could imply high risks in personal data processing. "Experts warn of the possible ways in which more and more processes and choices made by managers with regard to recruiting, remuneration and even dismissals are automated, too often giving free rein to discriminatory biases, perpetuating social segregation and impairing humanness and fairness"[11]. The HR manager is able to collect big data about the employee and is supported, if not replaced, in decision by algorithmic systems (the so called "boss ex machina")[12]. In digital transaction processes, algorithmic management can supervise, assign tasks, provide direct instructions by limiting the level of autonomy, and evaluate workers (including their performance, behaviour, earnings, and working conditions) up to and including dismissal.

Digital enterprises, that aim to measure their sustainability, have to estimate not only KPIs on the employee protection against sexual discrimination, the equal treatment in the company or the work and life balance, but also KPIs on healthy and safe work conditions and the real and transparent protection of personal data effectively recognized (according to article 22 GDPR and the directive EU no. 2019/1152).

---

11 Schubert, C., Hütt, M.T. (2019) 'Economy-on-demand and the fairness of algorithms', *European Labour Law Journal*, vol. 10(1).

12 Aloisi, A. (2022) *Rise of the boss ex machina: employer powers in workplaces governed by algorithms and artificial intelligence* in Lo Faro, A. (2022) *New Technologies and Labour Law. Selected topics*, Giappichelli.

Regarding a mere legislative compliance, the GDPR does not prohibit profiling or using algorithmic decision-making systems when these processes are necessary for entering into or performance the labour contract.

The EU Regulation obliges the employer, as Controller, to provide technical mechanisms that ensure the humanisation of the final decision, thus rebalancing the disproportion of contractual power in the employment field for employees exposed to automated algorithmic decisions[13]. The employer shall "implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision" see article 22(3) GDPR. Anyway in article 22(4) GDPR automated "decisions (…) shall not be based on special categories of personal data referred to in Article 9(1)". Referring to article 35(3) GDPR "a data protection impact assessment (…) shall in particular be required in the case of: a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person".

The General Data Protection Regulation does not ban automated decision or profiling process, but obliges companies, as data controller, to a fair data process according to a proportionality and data minimization, prior notice transparency and right to access (article5, 13 and 15(1)(h) GDPR)[14]. On the other hand, the employee, as data subject, has the right to obtain human intervention acts like a guarantee against the decisions of not-really-intelligent artificial intelligence (C71 and article 22(1) GDPR).

The EU regulatory technique of protection is to hold the employer, the data controller and to grant information rights to employees and trade union. This trend emerges clearly in the European Union digital strategy in the regulatory package still being proposed and has recently been codified in Italy in the transposition of EU Directive no.2019/1152. The Italian legislative decree no. 104/2022 introduces new EU rules on transparency and predictable work conditions enhancing employee's right and Personal Data Protection in HR digital

13 Lackovà, E. (2022) 'The Fragility of Pre-contractual Labour Relations in the Light of Algorithmic Recruitment', *Diritti Lavori Mercati*, vol. 2.

14 Controllers must provide meaningful information about the logic involved in the decision process, not necessarily a complex explanation of the algorithms used or the disclosure of the full source code, but a "sufficiently comprehensive explanation that allows the data subject to understand the reasons for the decision", Guidelines on Automated individual decision making and Profiling for the purposes of Regulation 2016/679" WP215 1 (2017). Transparency means that Controllers must provide to data subjects (e.g. workers) "relevant information related to fair processing, communicate and facilitate the exercise of their rights, enabling them to understand, and if necessary, challenge the data processing", Guidelines on Transparency under Regulation 2016/679" WP260 rev. 1 (2018).

process, only partially defined in article 22 and 88 GDPR. The Italian legislative Italian decree no. 104/2022 statues the employees' right to be informed on automated decision-making or monitoring systems and the duty to provide information to the union representatives and to the local/national level or to public bodies entities. According to article1 bis of Legislative Decree no. 152/1997 as amended by Italian Legislative Decree no. 104/2022, the employee has a special right to be informed on automated decision-making or monitoring systems. In Italy the private or public employer is required "to inform the worker of the use of automated decision-making or monitoring systems designed to provide relevant information for the purpose of hiring or conferring the task, managing or terminating the employment relationship, of the assignment of tasks or duties as well as indications affecting the supervision, evaluation, performance and fulfilment of the contractual obligations of the workers". Before the beginning of the labour relation, and in any case at least 24 hours before any change, in Italy the employee has the right to receive further information on the automated decision-making or monitoring systems regarding:

a. the aspects of the employment relationship affected by the use of the systems;
b. the purposes and purposes of the systems;
c. the logic and functioning of the systems;
d. data categories and key parameters used to program or train the systems, including performance evaluation mechanisms;
e. the control measures adopted for automated decisions, any correction processes and the person in charge of the quality management system;
f. the level of accuracy, robustness and cybersecurity of the systems and the metrics used to measure these parameters, as well as the potentially discriminatory impacts of the metrics themselves.

The employee, directly or through the company or local trade union representatives, has the right to access this information and to request further information. Employer is obliged to transmit the requested data and to give written answers within 30 days (article1 bis (3) Italian legislative decree n.152/1997 as amended by legislative decree n.104/2022).

The employer is required to integrate the information with the instructions for the worker regarding data security and the updating of the processing register concerning automated decision-making activities, including surveillance and monitoring activities. In order to verify that the tools used to carry out the work performance comply with GDPR provisions, the employer is obliged to carry out a risk analysis and an impact assessment of the same treatments, proceeding with prior consultation of the Guarantor for the protection of personal data where the conditions pursuant to article 36 GDPR exist. The information described must be communicated, in a structured format, commonly used and readable by automatic device, not only to the employee but also to

the union representatives at a local/national level or to the Italian Ministry of Labor and Social Policies (the Italian National Labor Inspectorate may request the communication of the same information and data). In Italy, accountability and transparency are the regulatory tools used to enhance the protection of employee in working relationship where automated decision-making and/or monitoring tools are used. Algorithmic management must be transparent and accountable for workers and trade union, since it has a significant impact on the work conditions effectively recognized to employees. "In the digital society, the limits of the individual dimension of judicial protection of interests affected by mass torts as well as the need and central relevance of collective judicial protection instruments in order to ensure effectiveness of rights, proper functioning of markets and respect for democratic values appear even more evident"[15]. The Italian legislative decree no. 104/2022 enforces the transparency rights on monitoring and predictable work conditions for employees and trade unions to reduce "algorithm opacity"[16]. That solution anticipates the EU Commission proposal on AI and Platform Work[17].

4. The sustainability of algorithmic management: transparency, fairness and accountability in the EU Commission proposal for artificial intelligence and work on platform

The legal basis of the proposal Artificial Intelligence Act COM (2021) 206 final 21st April 2021 liberalise the production and marketing of AI systems in the EU, provided that these systems comply with the standards of the Act.

15 Mazzei, G. (2023) 'Società digitale e collective redress: azione rappresentativa europea, class action statunitense e azione di classe italiana a confronto', Federalismi.it, vol.2. Available at: https://www.federalismi.it/nv14/articolo-documento.cfm?Artid=48302&content=Societ%C3%A0%2Bdigitale%2Be%2Bcollective%2Bredress%3A%2Bazione%2Brappresentativa%2Beuropea%2C%2Bclass%2Baction%2Bstatunitense%2Be%2Bazione%2Bdi%2Bclasse%2Bitaliana%2Ba%2Bconfronto&content_author=%3Cb%3EGabriella%2BMazzei%3C%2Fb%3E

16 "Algorithmic opacity can undermine the effective exercise of workers' rights for two reasons. On the one hand, workers may not realize that their rights have been violated. On the other, even if they do realize it, they may fail to acquire useful evidence to reveal in court the material truth behind the algorithms". Gaudio, G. (2022) 'L'algorithmic management e il problema della opacità algoritmica nel diritto oggi vigente e nella Proposta di Direttiva sul miglioramento delle condizioni dei lavoratori tramite piattaforma', Lavoro Diritti Europa, vol.1. Available at: https://www.lavorodirittieuropa.it/images/Gaudio_-_14_gennaio_2022_-_Articolo_LDE_-_032022.pdf.

17 Faioli, M. (2022) 'Trasparenza e monitoraggio digitale. Perché abbiamo smesso di capire la norma sociale europea', *Federalismi.it*, vol. 25. Available at: https://www.federalismi.it/nv14/articolo-documento.cfm?Artid=47826.

The AI Act adopts a preventive and proactive approach based on risk and its preventive categorization[18]. Degrees of risk warrant different regulation. Unacceptable risk systems use artificial intelligence in ways that threaten security and human rights and, consequently, are prohibited unless expressly exempted: subliminal techniques, exploiting people's unawareness or vulnerabilities (due to age or disability) to materially distort behaviour so as to cause harm (article5 Artificial Intelligence Act COM (2021) 206 final 21st April 2021). The high-risk category includes systems intended for use in the recruitment or selection of individuals (for advertising vacancies, screening or filtering applications, evaluating candidates in interviews or tests as well as making decisions on promotion and termination of contractual employment relationships – article 6 Artificial Intelligence Act COM (2021) 206 final 21st April 2021). In "high risk" systems, an integrated risk management system is adopted with compliance and information requirements (such as adequate technical documentation, data governance, and quality of training datasets, conformity assessment and declaration, registration and traceability, monitoring and supervision, measures appropriate human oversight, as well as requirements for transparency, accuracy, robustness and security), providing for controls and sanctions as enforcement tools (article 8 Artificial Intelligence Act COM(2021) 206 final 21st April 2021).

Finally, in low-risk systems there are specific transparency requirements while in minimal-risk systems free development and use are allowed.

In the AI Act, an anthropocentric artificial intelligence model takes shape in which various protection techniques are combined such as the principles of prevention, control and institutional cooperation[19]. These standards completely ignore the role of the social partners in regulating the introduction of technological tools at work. The current proposal presents a "techno-deterministic" approach to algorithmic monitoring and decision-making. In other words, digital managerial systems and practices are allowed in principle but the employer is responsible and required to assess the risk, take protective measures and inform the employee. In the Proposal for a Directive on platform work, it is planned to give platform workers enhanced information rights over and above those in the GDPR see article 13, 15 and 22[20]. A right to be informed about automated monitoring systems which are used to monitor, supervise or evaluate the work performance of platform workers through electronic means article 6, (1)(a), (2)

18 Finocchiaro, G. (2022) 'La proposta di Regolamento sull'intelligenza artificiale: il modello europeo basato sulla gestione del rischio', *Il diritto dell'informazione e dell'informatica*, vol. 2, pp. 303-322.

19 Alpa, G. (2021) 'Quale modello normativo europeo per l'intelligenza artificiale?', *Contratto e impresa*, vol. 4, p. 1018.

20 Proposal for a Directive of the European Parliament and of the Council on improving working conditions in platform COM (2021) 762 final 2021/0414 (COD) Brussels, 9.12.2021. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:52021PC0762.

(a). A right to be informed about automated decision-making systems which are used to take or support decisions that significantly affect those platform workers' working conditions article 6, (1)(b), (2)(b). A right to receive a written explanation about how the automated decisions are reached and to access a competent human contact to discuss and to clarify the facts, circumstances and reasons leading to a decision, asking the platform to review a detrimental decision article 8. Moreover, the Proposed Directive provides that, in disputes concerning the qualification of an employment relationship with a digital platform, "national courts [...] may order the digital employment platform to disclose any relevant evidence within its control" article16 (1).

In the proposed Directive, each member State shall limit the employer's authority to process workers personal data. The proposal bans some of the most abusive forms of data processing, including on "any personal data on the emotional or psychological state" of platform workers, data concerning their health, and private conversations and collecting "any personal data while the platform worker is not offering or performing platform work" (see article 6 par. 5)[21]. According to article 7, digital labour platforms will have to evaluate the risks of automated monitoring and decision-making systems to the safety and health of platform workers and ensure that such systems do not in any manner put undue pressure on platform workers or otherwise put at risk the physical and mental health of platform workers. Moreover, it is requested to digital labour platforms to entrust competent people the monitoring automated systems and to protect platform workers from negative consequences (such as dismissal or other sanctions) for overriding automated decisions. Last but not least in the Directive importance is given to union protection tools (article 14).

In the European regulatory framework, algorithmic transparency is accompanied by the necessary transparency on the part of those who govern the algorithms (in EU Regulation no. 2022/2065 Digital Services Act on the platforms bear the duties of due diligence and assessment of systemic risks related to the use of algorithms, transparency obligations and motivation towards users) and the class action tools granted to users. In the European Commission's proposed Regulations and Directives, monitoring, tracking, and automated decisions are considered high-risk activities. It will be crucial define social taxonomy

21 In particular the platform shall not process any personal data on the emotional or psychological state of the platform worker; shall not process any personal data relating to the health of the platform worker, except in cases referred to in article 9(2) points (b) to (j) GDPR; shall not process any personal data in relation to private conversations, including exchanges with platform workers' representatives; shall not collect any personal data while the platform worker is not offering or performing platform work article 6(5). De Stefano, V. (2022) 'The EU Commission's proposal for a Directive on Platform Work: an overview', *Italian Labour Law e-Journal*, vol. 15(1). Available at: https://doi.org/10.6092/issn.1561-8048/15233.

sustainability social standard for companies involved in digital transformation, as digital transformation could lead to high risks in the processing of personal data.

5. Promoting corporate sustainability through the corporate disclosure

The human-centric model of artificial intelligence governance, chosen at the european regulatory level, highlights how social sustainability implies a conscious and socially responsible corporate use of data and algorithms.

Through the Corporate Sustainability Reporting Directive (CSRD) companies will be required to make disclosures about the role of the governance bodies, management and control of sustainability will be required to introduce forms of incentives for governance members, related to the achievement of the same goals and will be required to report counting on the allocation mechanisms.

Companies will have to provide sustainability information regarding the impact of their activities on people and the environment (inside-out approach), as well as on how sustainability factors affect them and their results (outside-in approach). In reporting sustainability initiatives, not only financial statement information should be considered, but also material impacts, risks and opportunities related to the entire upstream (upstream) and downstream (downstream) value chain must be considered as well as the results of due diligence activities (as also indicated in the proposal of the new Corporate Sustainability Due Diligence Directive) and materiality analysis.

It will remain mandatory for corporate sustainability to determine how to reconcile such guarantees with the use of machine and deep learning systems, which are capable of self-learning and thus operate according to decision-making processes that are opaque and unpredictable to the programmers themselves.

On 14th February 2024 the Council and the European Parliament reached a provisional deal on a directive on time limits for the adoption of sustainability reporting standards for certain sectors and for certain third-country companies, amending the Corporate Sustainability Reporting Directive (CSRD). The agreement will give companies more time to prepare for the sectorial European Sustainability Reporting Standards (ESRS) and for specific standards for large non-EU companies, which will be adopted in June 2026, two years later than the originally scheduled date. This was a major mistake. Only a real mandatory disclosure will allow access to human oversight and eventual correction of the system, enabling sustainable protection of the individual through the remedies of data protection, non-discrimination and labor law.

Chapter III
Metaverse and "Meta" Crimes, Are We Facing New Threats for People Rights?

by Emilio Sacchi*

INDEX: 1. What is the metaverse? – 2. How many types of crimes can be perpetrated in the metaverse? – 3. How can we prevent and prosecute meta-crimes?

1. What is the metaverse?

The first step to solve a problem is to acknowledge its existence.

So, to make the metaverse a better place it is necessary to show its importance in our future. To do so, first we need to understand what the "metaverse" is and what its boundaries are and will be in our lives. Usually, people think that metaverse is a new virtual reality where you enter only voluntarily to entertain and escape from the "real world" and that is a place built for nerds and weirdos where they can play and do their tech stuff. But they are mistaken.

The first goal to reach a higher level of safety in the metaverse is to make people understand that it is a new dimension of reality (which could be augmented or fully virtual) and sooner or later it will concern everyone, just like internet did in the first place, social followed, and so on.

The term "metaverse" was first used by Neal Stephenson in his cyberpunk book "Snow crash" in the 1992 and the Treccani Encyclopedia define it as a "three-dimensional space where people can move, share and interact through avatars"[1]. So, one of the major goal the metaverse aims to achieve is to use technology to create a new immersive reality where people can live, maybe, forgetting to be in it. This is something called illusion of nonmediation[2]. The more the technology provides the stimulation that human brain expects, the higher

---

* Criminal Lawyer, co-founder Networklex Studio legale associato, cybersecurity and cybercrime expert, auditor iso 42001:2023.

1 See: https://www.treccani.it/enciclopedia/metaverso_(Lessico-del-XXI-Secolo)/.

2 "It occurs when a person fails to perceive or acknowledge the existence of a medium in his/her communication environment and responds as he/she would if the medium were not there. Although in one sense all of our experiences are mediated by our intrapersonal sensory and perceptual systems, "nonmediated" here is defined as experienced without human-made technology". Lombard, M., Theresa Ditton, T. (1997) 'At the Heart of It All: The Concept of

the emotional involvement and physical present of the subject is[3]. That said, is true that the metaverse is strictly connected to web 3.0, crypto, blockchain and NFT but not exclusively.

Indeed, if we put aside for a while the most popular metaverses like decentraland, sandbox, heroverse and so on and we focus for some low level/high efficiency futures that come with this new technology, we realize that there must be dozens of metaverses.

For instance, there are now many virtual private meeting rooms that are – as a matter of fact – a kind of metaverse that require a strict regulation because each participant will have an avatar that can interact with one another not only through voice but also with non-verbal communication. This kind of "small" virtual reality requires a regulation as well as the big and famous metaverses because similar actions committed in both types of metaverses may be considered crimes. The most popular example of a crime committed in the metaverse – precisely in the Meta's Metaverse called Horizon Worlds – was the sexual harassment reported by Nina Jane Patel in May 2022[4]. That event, even though it would have been perpetrated completely in virtual reality, compelled Zuckerberg to make some algorithmic adjustment to prevent it from happening again. Indeed, he introduced a protective bubble that block anyone from interacting with the avatar but also prevents the user itself from enjoying any content in the virtual reality[5].

The Author is aware that nowadays metaverses are quite empty but he thinks that in the years to come they will be crowed places where everybody has to be. Gartner Inc predicts 25% of people will spend at least one hour per day in the metaverse by 2026[6]. We need to understand that the evolution of technology goes that way and the world needs to adapt to it. Major player in the world is already set and ready for the new metaverse experience and jobs will follow that path. For instance, marketing, public relations, sponsorship in this new reality will create new jobs which will increase the flow of economic and financial resources.

Presence', Journal of Computer-Mediated Communication, vol. 3(2). Available at: https://doi.org/10.1111/j.1083-6101.1997.tb00072.x.

3 Ingarrica, D. (2022) 'Metaverso criminale. Quali interazioni nel presente nazionale e quali sfide globali del prossimo futuro', Giurisprudenza Penale Web, vol. 9. Available at: https://www.giurisprudenzapenale.com/2022/09/05/metaverso-criminale-quali-interazioni-nel-presente-nazionale-e-quali-sfide-globali-del-prossimo-futuro/.

4 See: http://repubblica.it/esteri/2022/02/14/news/metaverso-337711044/.

5 See:https://www.altalex.com/documents/news/2022/02/23/il-metaverso-e-il-reato-di-molestie-sessuali-nella-realta-virtuale.

6 See: https://www.gartner.com/en/newsroom/press-releases/2022-02-07-gartner-predicts-25-percent-of-people-will-spend-at-least-one-hour-per-day-in-the-metaverse-by-2026.

2. How many types of crimes can be perpetrated in the metaverse?

Once we cleared what is a metaverse and where we can encounter it, we need to find out what kind of crimes we can face in it. I think that an avatar can commit or be the target of three types of crimes. The first one is a criminal activity that is perpetrated through the metaverse but has effects only in the real life, such as a threat uttered by an avatar to a real person. In this case, it does not matter how it was made or where, because for our legal system there is no difference between a social network or a virtual reality when they are used as tools to commit a crime. An example of this is a threat of violence towards a person through an avatar or cyberbullying.

The second type is a felony that is committed in the metaverse and has effects both in real life and in virtual one. For example, a fraud committed in the metaverse and involving goods that have an economic value (such as an NFT) produces a double effect. On one hand, the target of the scam loses the money (or more likely the cryptocurrency) he used to purchase the fake NFT (this is the "real life effect of the crime) and, on the other hand, he/she is not able to "use" the NFT in the metaverse (this is the virtual life damage caused by the crime). Another example is the sexual harassment suffered by Nina Jane Patel, I've mentioned earlier. In such cases there has to be a way to pursue the criminal and protect the victim's rights both in the real life and in the virtual one.

Last but not least, it's a felony that is committed in the metaverse and has effects only there. This would be the case of private violence or stalking perpetrated from an avatar to another one. These crimes do not require physical misconduct to be committed and can be perpetrated in the metaverse. For example, an avatar can follow another one all day long, making it feel unsafe and not free to act normally in the metaverse, just like a "normal" stalker.

Another type of crime that can be committed in the metaverse is the impersonation or identity theft, which occurs when someone falsely represents themselves as another avatar to gain an advantage. In those cases, even if nothing seems to happen in the real word, it's necessary to ensure protection for the victims of those crimes and bring criminals to justice because the victims' rights can be infringed due to the strict connection between a person and their avatar. It needs to be understood that the fact that an avatar is not a real person does not prevent it from being a target of a crime because it carries certain rights that need to be protected. As we said before, there could be a strong connection between the person and the avatar that may cause a real suffering and a real danger for someone's right even if it happened only in the virtual world.

We can all agree that this kind of threats would be very rare right away due to the lack of interest for the metaverse among the most of the world's population, but things may change in the future, and we need to be prepared. We

cannot afford to introduce people (who are likely minors) to a non-regulated environment where they can talk, act, maybe "work" and – in some way – live without any rules or without any supervision by a moderator or someone who can eventually intervene in dangerous situation created among avatars. Although it is important to establish a balance between rights and duties because we all know that too many rules mean fewer participants because of the bad user experience. Imagine a huge metaverse full of features to interact with other avatars or NPCs but also with so many rules which are hard to understand and even harder to follow.

It would be a disaster and all the users will move to a freer (and dangerous) metaverse.

3. How can we prevent and prosecute meta-crimes?

We cannot accept that the only way out of a dangerous situation in the metaverse is to log out of it or move to a safe zone we mentioned before, because that is not the exercise of a "user behind the avatar" right but likely the effect of being targeted by a criminal.

The Author thinks we need more than that, like a multi-level security system based on the threat faced by the avatar and using the AI to process the information as quickly as possible to give the assistance the user needs. We probably also need an authority that can report bad behavior in metaverse outside of it, in the real world, to ensure that any illegal action will be punished. To do so we face two major problems: how to recognize a criminal activity and how to punish its author. These two questions lead us to another technical issue: do we already possess the tools to solve these two problems? Do we need an international metaverse law? How can it work with all the problem with jurisdiction and rights among the whole world? The Author thinks it's too early to have the right answers but not too early to ask ourselves these questions.

For instance, the Author thinks that – at least in Italy – to convict someone for a felony committed in the metaverse we should look at what we do regarding crimes perpetrated on social media and so the competent court is determined based on where the legally protected right was violated.

The AI may help us to recognize possible threats to avatars in the metaverse, monitoring both language and non-verbal communication. Once the AI recognize a potential threat, it may suggest the users/avatars involved to stand down and, if they don't, the AI may give the opportunity to the targeted avatar to request immediate support from a moderator or to record what is happening and submit it to the host of the metaverse. Otherwise, the AI may have the ability to freeze the avatar responsible for the threat and investigate on it.

What about punishment? How can we reach someone who could be 10.000 kilometers away from us? I think we may focus on a system that can convert jail

time in fine or ban the "criminal" user for a certain amount of time, in every metaverse. This would reach both rehabilitative and preventive purpose of the punishment. It is also possible to determine that who violated multiple time another user right could be described as violent or abuser to warn other avatars. However, these solutions can lead to two other problems: a) how to identify the author of a misconduct; b) how to determine when someone earn the status of abuser/violent.

To solve these problems, the Author thinks that we need to aim for a central authority – hypothetically a private blockchain – were everyone need to be registered to got an unique asymmetric couple of keys mandatory to interact with every kind of metaverses, like a passport to get a virtual avatar. In this way it can be possible either to identify the author of a crime and to punish them with a fine or even a ban.

The first one could be executed forbidding any kind of payment towards a certain wallet in the metaverse, redirecting any kind of transaction to another public key and use that amount of "money" for charity purposes. The ban, instead, could be executed forbidding a private key to log in into any metaverses for a certain amount of time (hours, days, months, years). Of course, all the users need to be aware of the rules and to ensure that knowledge it's even possible to introduce a random quiz that anyone has to clear to enter in a metaverse.

The Author is aware that the speech raises more questions than those it answers, but he thinks that's the point of study the involvement of new technologies in people lives.

Only if you make yourself the right question you may be able to solve the problem referred to it. The best is yet to come.

Chapter IV
Harmful Contents Online and Platforms Criminal Responsibility

by Beatrice Panattoni*

INDEX: 1. The context. – 2. Content-related offences. – 3. Online platforms responsibility. – 4. The contribution of criminal law to the debate on tech regulation policies.

1. The context

Life around the world is increasingly mediated by digital platforms. Individual experiences, manifestations of one's personality, and intimacy take shape through the continuous uploading and systematic sharing of data online, where bodies become screens and lives become software code. We are, in philosopher Luciano Floridi's words, "informational beings"[1].

This transition to a digital age has led to the emergence of new criminal phenomena based on both harmful contents and abuses of lawful contents, which are not always easily traceable to existing criminal offences. Content-related offences take place within and through digital communication services (social media), which have been left, until recently, without adequate legal regulation. The range of criminal offences referable to this area is significantly broad: it includes offences against public interests (such as hate speech, incitement or apology to crime and violence, dissemination of terrorist contents), and offences against a specific victim (such as forms of interpersonal hatred, non-consensual pornography, child pornography).

The scale of the dissemination of harmful contents online is quantitatively high and qualitatively severe. In many cases, they generate irreversible consequences. The technical architecture of cyberspace and content-sharing platforms (social platforms) greatly amplifies the scope and consequences of

---

* Postdoctoral researcher in Criminal Law at the University of Verona, Italy. She earned her Ph.D. in Criminal Law from the same University in 2022. She has been visiting scholar at the University of Freiburg (Germany), at the Max Planck Institute for the Study of Crime, Security and Law, and at the University of Washington (US). Her research investigates how digital technologies challenge theories of criminal responsibility attribution, focusing primarily on online platforms responsibility and AI-related crime.

1 Floridi, L. (2014) *The Fourth Revolution: How the Infosphere is Reshaping Human Reality*, Oxford.

communicative conducts[2]. From a quantitative point of view, there is a fast duplication of the same or similar content, which can reach a potentially indeterminate number of recipients. From a qualitative point of view, access to and sharing of content is increasingly immediate and within everyone's reach. Given that the scale of dissemination is directly linked with both the seriousness and the frequency of the harms that might be caused by illegal contents online, tech policies could arguably require new criminal policies aimed at the prevention of such crimes[3].

To develop effective measures and policies aimed at preventing content-related offences, the role of private actors which manage the social platforms where these crimes are realized is a necessary step. Therefore, an evaluation of the adequacy and efficiency of tech policies on online platform responsibility, assessing whether criminal policies, alongside other measures in place, should be included.

2. Content-related offences

A categorization of the relevant offences within the category of "content-related offence" allows giving homogeneity and systematicity to the criminal phenomenon of unlawful contents online. The choice to base the categorization on the nature and the seriousness of the harm aims to provide a standard of care to the actors liable for the removal of that content, which must be mitigated by a case-by-case analysis.

Given the breadth of the offences that might be realized through communicative conducts online, certain differentiations should be made. Specifically, we can suggest dividing content-related offences into two main clusters, based on the nature of the harm they create. The first cluster includes offences that harm public interests, such as public order or human dignity. The second cluster includes offences that harm the individual rights of a specific victim, such as her/his reputation or sexual freedom. The paper will consider, as case studies for each cluster, hate speech online on the one hand, and gender-based cyber-violence[4] on the other hand.

---

2 Among studies of psychology see Aiken, M. (2017) *The Cyber Effect*, New York.

3 Digital technologies represent a facilitator in the shift of criminal law from been reactive to been based on crime control and risk avoidance. See Koops, B. (2009) 'Technology and the Crime Society: Rethinking Legal Protection', *Law, Innovation & Technology*, vol. 1(1), pp. 93-124, ILT Law & Technology Working Paper No. 010/2009.

4 Among reports addressing the topic of "cyber violence against women and girls", see: report of Sept. 24, 2015, drafted by the Broadband Commission for Digital Development, a body established in 2010 by the International Telecommunication Union (ITU) and the United Nations Educational, Scientific and Cultural Organization (UNESCO) at the behest of the then secretary-general of the United Nations; recommendation No. 35 of July 26, 2017, aimed at updating the previous Recommendation No. 19 of 1992447, prepared by the United

The circulation of violent content based on hatred toward certain minorities has exponentially increased in recent years, which was also worsened during the COVID-19 pandemic[5]. As stated by the European Agency for the Protection of Fundamental Rights (FRA), "Online hate has taken root in European societies"[6]. Given this framework, European institutions have begun to address the updating of legal sources on the subject. The definition of hate crimes and hate speech is harmonized by the Council Framework Decision 2008/913/JHA of November 28, 2008, on Combating Certain Forms and Expressions of Racism and Xenophobia by means of Criminal Law, in which manifestations of hatred based on race, colour, religion, ancestry, or national or ethnic origin are covered. However, this definition is limited, if we consider the European Convention and Charter of Human Rights (article 14 ECHR, article 21 EUCFR), where the prohibition of discrimination is extended beyond those based on racial or xenophobic grounds. The most notable gap concerns discrimination based on gender, sex, and sexual orientation.

Hate speech online is particularly affected by the technical architecture of online platforms, deepening the harm it can cause. The use of algorithmic agents that profile users can lead to opinion polarization, inducing the user to view hateful materials on a loop, as they are qualified by the algorithm as "similar" to those usually consulted by the user. Amplifying the resonance of hate content that glorifies violence creates a higher possibility that words become violent actions[7].

Given its sudden explosion in the digital age, the need to update a European-level harmonization of the legal discipline related to hate speech, as well as its definition, has led the Commission to issue a Communication to the European Parliament and the Council with the aim of triggering a Council decision identifying hate speech and hate crimes as areas of serious and transnational crime, which meet the criteria specified in the first subparagraph of article 83(1) of the TFEU, so that substantive legislation harmonizing the definition and penalties for hate speech and hate crimes can subsequently be proposed directly

Nations Committee on the Elimination of All Forms of Discrimination against Women (CEDAW); the United Nations Special Rapporteur's Report of June 18, 2018 on Online Violence.

5 Numerous statistical data are reported by Peršak, N. (2022) 'Criminalising Hate Crime and Hate Speech at EU Level: Extending the List of Eurocrimes Under Article 83(1) TFEU', Criminal Law Forum, vol. 33, pp. 85-119. Available at: https://link.springer.com/article/10.1007/s10609-022-09440-w.

6 See FRA, Overview of antisemitic incidents recorded in the European Union 2009-2019, Publication Office of the European Union, 2020.

7 See Forti, G., Lamanuzzi, M. (2022) *Digital Violence: A Threat to Human Dignity, a Challenge to Law*, in D. E. Vigan, E., Zamagn, S., M. S. Sorond M.S. (eds.) (2022) *Changing Media in a Changing World*, Città del Vaticano, vol. 183.

by the EU[8]. Regarding this legislative initiative, the qualification of hate speech as a serious and transnational crime will have to be addressed based on the harm principle, considering it, however, from a human-centric perspective, as a crime that violates human dignity rather than the public order. A similar dramatic evolution also characterizes gender-based cyber-violence, which has many forms. They include: "online sexual and psychological harassment, cyberbullying, online stalking, non-consensual pornography, online sexist hate speech, and new forms of online harassment such as zoom bombing or online threats"[9]. As pointed out by the European Parliament, online forms of violence disproportionately affect women and girls and must be understood as an inseparable continuum from offline violence, as both are interconnected.

European institutions have begun to outline legislative policies in this area as well, which became complementary to those dedicated to countering online hate. Gender-based violence has also been identified as an area of crime that meets the criteria of article 83(1) TFEU, and the European Parliament has asked the Commission to submit a proposal for a Council decision identifying gender-based violence as a new serious and transnational crime, then using it as the legal basis for a holistic, victim-centred directive aimed at preventing and combating all forms of gender-based violence, both online and offline[10].

The seriousness of specific forms of cyber-violence should not be underestimated just because these criminal conducts are perpetrated online. Violence in the digital dimension, where the victim's "physical" body is not directly involved, may lead to a flawed criminal framing of some events, especially those involving the victim's sexual sphere. In the field of sexual abuse through images (or non-consensual pornography)[11], according to the criminological theory of the so-called "embodied harm"[12], the body of the victim is equally involved in

8 Communication from the Commission to the European Parliament and the Council, A more inclusive and protective Europe: extending the list of EU crimes to hate speech and hate crime, 9.12.2021 COM (2021) 777 final. For a comment on the proposal see Peršak, N. (2022), 'Criminalising Hate Crime and Hate Speech at EU Level: Extending the List of Eurocrimes Under Article 83(1) TFEU', *Criminal Law Forum*, vol. 33, pp. 85-119. Available at: https://link.springer.com/article/10.1007/s10609-022-09440-w.

9 European Parliament resolution of 16 September 2021 with recommendations to the Commission on identifying gender-based violence as a new area of crime available at: https://oeil.secure.europarl.europa.eu/oeil/cs/procedure-file?reference=2021/2035(INL)#gateway.

10 Ibid.

11 See Citron, D., Franks, M. (2014) 'Criminalizing revenge porn', *Wake Forest Law Review*, vol. 49, p. 345, (for the US); Gillespie, A. (2015) "Trust me, it's only for me': 'revenge porn' and the criminal law', *Criminal Law Review*, vol. 866, (for UK); and, among Italian scholars, Caletti, G. (2019) 'Libertà e riservatezza sessuale all'epoca di internet. L'article 612-ter c.p. e l'incriminazione della pornografia non consensuale', *Rivista italiana di diritto e procedura penale*, Anno LXII, Fasc. 4.

12 See Powell, A., Henry, N. (2017) *Sexual Violence in a Digital Age*, London. It elaborates, by applying sociological studies of "embodiment" (which conceive of the "body" not as mere

the criminal act. Therefore, the right harmed by these behaviours should not be reduced to confidentiality alone, whereas it should include the freedom of sexual self-determination as well.

Besides policies aimed at the direct criminalization of hate speech and gender-based cyber-violence, the complexity of the digital context requires and justifies criminal policies aimed not only at combating such crimes but also at preventing them. In addition, then, to hold responsible the users, tech policies must also look at the role and responsibilities of social platforms.

Content-related offences have been qualified also as "platform-enabled crimes"[13], since, in many cases, their realization is enabled, facilitated, or amplified by the architecture of services offered by online platforms. Indeed, the harmfulness of unlawful contents online does not end with the uploading, it only starts from there. Recognizing legal relevance to the entire lifecycle of online information aims at understanding and analysing the criminal phenomenon in its entirety. It is precisely the persistent availability of harmful content that causes the most incisive consequences for the victim. It is from a single upload that a chain effect of sharing and further dissemination, which represent the core of the potential harmfulness of such behaviours, is most often triggered. Thus, it is what happens after the uploading that constitutes one of the most incisive innovations of the digital dimension. These stages cannot be considered legally irrelevant, they should instead be regulated.

This new scenario has led to the emergence of a protection gap in the EU and Member States' legal systems, leaving a sector of activity as fundamental to social life as digital services without an adequate legal framework. To fill this considerable gap in legal protection, the EU institutions are following a two-way legislative strategy, aimed, on the one hand, at combating and criminalizing the dissemination of offensive materials online, and, on the other hand, at regulating the digital communication services sector by providing an apparatus of specific legal obligations to online platforms. These two policies should be understood as complementary to each other.

The operators of digital services where illicit materials circulate are the only entities able, not only to actively intervene after the information is placed online, but also to design the technical architecture that makes possible (or determines, as the case may be) the uncontrolled increase of violent, harmful, or dangerous content. Therefore, the responsibility of online platforms should not be limited to content moderation, whereas it should consider the responsibility related to the online platforms' use of autonomous algorithmic agents in managing their

physical integrity, but as a physical, metaphysical, social and cultural phenomenon) to digital experiences, the concept of "digital-embodied harm", according to which it is necessary to engage the corporeality of victims in the experience of cyber-violence episodes.

13 See Hamilton, R.J. (2022) 'Platform-Enabled Crimes: Pluralizing Accountability When Social Media Companies Enable Perpatrators To Commit Atrocities', *Boston College Law Review*.

services, and, more specifically, in indexing, filtering, and obscuring content, as provided in their internal policies.

3. Online platforms responsibility

Given that content-related offences are criminally relevant, it remains to be established whether online platforms (also defined as hosting providers)[14] might be held criminally responsible for the content-related offences committed by their users.

Among Italian scholars[15], it is controversial whether a platform can be held criminally liable in case of illegal contents hosted on its service. Since crimes realized by uploading illegal content online are usually considered normatively "concluded" at the moment of the "publication" of the content, no criminal liability is configurable after that moment. However, thanks to the use of new technologies and new tools (such as algorithms capable of filtering and indexing), new scenarios are opened. Starting from the category of "active" hosting provider, elaborated by the European Court of Justice[16], it has become clear that the new technologies implemented by hosting providers allow a higher and stricter regime of liability, making them sort of "publishers" of the contents hosted on their platforms.

However, resorting to criminal law in this context presents several shortcomings. The technical architecture of the network and of digital communication services have an impact not only on the crimes' harmfulness, but also on the subjective element of "intent". If what happens after the upload of the content online (i.e., the persistent and uncontrolled circulation of the same or similar content), becomes legally relevant, it cannot in any case be considered as represented and intended by the original user who made the original upload, except resorting to forms of strict liability.

The same can be said for online platforms, whose intent is difficult to ascertain. Indeed, it would be necessary to establish that the platform's manager had actual knowledge of the illicit content and that it intentionally failed to remove it. If we consider the large social platforms and the amount of information,

14 Hosting provider is the broader category which also include platforms. They are defined by the Digital Services Act (DSA), Regulation (EU) 2022/2065, as providers of services "consisting of the storage of information provided by, and at the request of, a recipient of the service". See article 3 lett. g) DSA.

15 Among the most recent contributions by Italian scholars see Fiorinelli, G. (2022) 'L'attuale ruolo del provider nella società digitale: modelli di responsabilità penale', *La Legislazione Penale*; Lamanuzzi, M. (2021) 'Il 'lato oscuro della rete'. Odio e pornografia non consensuale. Ruolo e responsabilità dei gestori delle piattaforme social oltre la net neutrality', *La Legislazione Penale*; Braschi, S. (2020) 'Social media e responsabilità penale dell'Internet Service Provider', *Rivista di diritto dei media*.

16 The leading case is European Court of Justice, judgment of March 23rd 2010, C-236/08 a C-238/08, Google France e Google, EU:C:2010:159.

they manage every second, it is not always possible to ascertain a malicious culpable behaviour for which they can be criminally liable. The scenario becomes even more complicated if we consider the autonomous "actions" of algorithms that "highlight" illicit content. Even for the automatic and autonomous functioning of such algorithmic "agents"[17], the platform operators cannot be said to have acted with criminal intent, and therefore they cannot be considered guilty.

4. The contribution of criminal law to the debate on tech regulation policies

In the area of unlawful content online, criminal policy choices can follow two main directions, depending on the legal system taken into consideration:

1. holding criminally liable the individuals involved (e.g., the managers of the platforms);
2. holding criminally liable the corporation, also through pecuniary sanctions, seen as expressions of the broader category of "punitive law".

The first option will present the same shortcomings that have been briefly outlined. Moreover, we need to keep in mind the specific features of the context where crimes related to digital technologies occur. Specifically, one of the main challenges of crimes related to digital technologies is that ICT's automatic and autonomous functioning "contaminates"[18] the manifestation of the *actus reus* and, consequently, they may influence also the *mens rea* requirement. This conceptual shift determines the impossibility of asking a single human agent to completely "control" what happens on digital services. Thus, in a context of "distributed moral responsibility"[19], where identifying the specific sources of responsibility becomes more difficult, resorting only to individual responsibility models based on the capacity to "control" harmful outcomes is not a suitable policy option.

In the digital society, resorting only to the "traditional" concept of individual responsibility based on the capability of control over one's own actions and their outcomes might become a dangerous policy choice. The prevention and contrast of harms related to digital technologies that fall only on citizens'

17 On the harms related to AI applications see Abbott, R., Sarch, A. (2019) 'Punishing Artificial Intelligence: Legal Fiction or Science Fiction, in University of California', *53 UC Davis Law Review*; Beck, S. (2016) 'Intelligent agents and criminal law – Negligence, diffusion of liability and electronic personhood', *Robotics and Autonomous Systems,* vol.86, pp. 138-143.

18 This suggestion is extracted from Bruno Latour's Actor-Network Theory (ANT), which suggests that in contemporary society the action is the result of the cooperation of multiple agencies, human and non-human. Therefore, we need to shift our attention from the subject and his/her actions to the interactions that bond the different entities (persons, technology, organizations, etc.) that create a reality of network. See Latour, B. (2007) *Reassembling the Social: An Introduction to Actor-Network-Theory*, Oxford.

19 See Floridi, L., Taddeo, M. (2016) 'The Debate on the Moral Responsibilities of Online Service Providers', *Science and Engineering Ethics,* vol. 22(6), pp. 1575-1603.

obligations might indeed create a slippery course that does not take into account the social transformation brought about by digital technologies.

Given the described shortcomings, criminal law might aim at its self-containment in this area of regulation, leaving it only to alternative measures, such as administrative sanctions. However, excluding criminal law completely from the set of policies could miss the opportunity to take advantage of the affirmative contribution it could bring in shaping responsibility models through its principles and guarantees, first and foremost the principle of culpability, without which these policies could become mere formal orders to be complied with, lacking an impact in terms of prevention and re-education.

Therefore, the second policy option seems more suitable. Grounding responsibility not on single persons but on corporations could provide more effective legal protection, as well as increase trust in digital technologies by creating a framework based on the proactive cooperation between the different stakeholders involved (the public, private actors, and authorities). Therefore, accountability models based on punitive law[20], which mitigate the "paradigm of control", might bring a significant contribution to the definition of the set of tech policies on platform-enabled crimes.

Despite which path will be chosen by national legislators, tech policies should avoid holding platforms indirectly responsible for the crime of their users, whereas they should hold them directly accountable for the guilty management of their service. This is also the approach followed by the EU Regulation on digital services (DSA)[21], which regulates several due diligence obligations, not aimed at preventing an unlawful event, but at setting a standard of care that is both adequate and socially accepted for large platforms, based on a virtuous organization of their services and businesses[22].

20 Wrongdoings punished with administrative sanctions that, due to the "suffering" they impose, preserve a punitive nature. On the topic see Dyson, M., B. Vogel, B. (eds.)(2018) *The limits of criminal law*, Cambridge.

21 See footnote n. 14.

22 See Montagnani, M.L. (2020) *A new liability regime for illegal content in the digital single market strategy,* in G. Frosio (eds.), *Oxford Handbook of Online Intermediary Liability,* Oxford, p. 399. The DSA provides for different sets of positive obligations for hosting providers, which increase in number and complexity according to the type of provider. The most stringent set of obligations concerns large platforms (with an average monthly number of active users of 45 million or more). These private actors will be obligated to set up a system for assessing and managing the "systemic risks" associated with their services, including the risk of dissemination of illegal content through their services (article 33 DSA).

PART II
Privacy, Data Protection and Data Governance

Chapter V
The Lies and the Fights for Privacy: Protecting Privacy and Human Dignity in the Digital Age

by Elena Pagani*

INDEX: 1. Introduction. – 2. Article 8 of the European Convention for the Protection of Human Rights and Fundamental Freedoms. – 3. Conclusion.

1. Introduction

The rapid development of digital technologies is improving our daily lives, while raising concerns about privacy and human dignity.

New technologies have introduced new privacy violations, including revenge porn, blackmail, and online hate speech, which pose significant risks to both online and offline communications. Such types of abuse can cause serious harm to individuals, particularly young women and marginalised groups, who are often targeted more frequently and more severely than other groups.

New forms of online gender-based violence, facilitated by the use of information and communication technologies, are a global concern. Rapid digital and technological developments, including the implementation of Artificial Intelligence and the Internet of Things, will inevitably give rise to new and different forms of cyber violence against women, with serious individual and collective implications worldwide[1].

Everyday women's online experiences differ from the ones of men. According to a 2016 survey, 53% of women experienced harassing behaviors versus 40% of men in the U.S.[2]. As reported by a study conducted for Amnesty

---

* Elena Pagani received his master's degree in Law with a thesis in legal information technology – University of Milan. She is currently a lawyer at Perani Pozzi Associati boutique law firm in Milan, dealing with privacy, personal data protection and advising on civil law issues, primarily in the area of new technology law. She is the author of contributions ranging from issues related to the right to be forgotten and the right to disconnect.

1 Stringhi, E. (2020) *Revenge Porn On Online Platforms: Legal Interpretations And Approaches To Combat Non-Consensual Intimate Image Distribution.* Privacy e innovazione [Master Law Thesis], vol. 6, Università degli Studi di Milano.

2 Data and Society Research Institute, (2016). *Online Harassment, Digital Abuse, and Cyberstalking in America*, New York.

International, 23% of women has experienced abuse or harassment online on one or more occasion, with significant social and psychological consequences[3].

In 2017, a survey by the Pew Research Center in the United States, revealed that while men experience slightly higher levels of online harassment than women, such as name-calling and physical threats, women are much more likely to experience severe types of gender-based or sexual harassment: 21% of women aged 18 to 29 reported being sexually harassed online, more than twice the percentage of men in the same age group (9%)[4].

An EU survey conducted in 2014 found that 1 in 10 women in the European Union report having experienced cyber-harassment since the age of 15 (including having received unwanted, offensive sexually explicit emails or SMS messages, or offensive, inappropriate advances on social networking sites). The same research suggests that up to 90% of "revenge porn" victims are female and that this number is increasing. The impact and harms caused by online abuse are also disproportionate[5].

The Revenge Porn Helpline – part of the UK Safer Internet Centre (UK SIC) – says that the number of reports of revenge porn has gone "through the roof" during lockdown[6]. The helpline, which was set up in 2015, saw a 98% increase in cases in April 2020, compared with April 2019. In August 2020, the helpline dealt with 285 cases, a 63% increase on the 175 dealt with in August 2019. In 2020, the hotline has dealt with 1,914 reports of revenge porn.

Women are disproportionately the targets of certain forms of cyberviolence compared to men.

There are a range of different terminologies used to describe this phenomenon, including gendered cyberhate, technology-facilitated violence, tech-related violence, online abuse, hate speech online, digital violence, networked harassment, cyberbullying, cyberharassment, online violence against women, and online misogyny[7]. Online misogyny, or its preferable synonym cybermisogyny, is "an umbrella term for all kinds of negative experiences that women can go through online because of their gender"[8].

---

3 Mori, I. (2017) 'Online abuse and harassment', *Ipsos*. Available at: https://www.ipsos.com/en-uk/online-abuse-and-harassment.

4 Ging, D., Siapera, E. (2018) 'Special issue on online misogyny', *Feminist Media Studies*, vol. 18(4), pp. 515-524; Pew Report on Online Harassment. Available at: http://www.pewinternet.org/2017/07/11/online-harassment-2017/.

5 Ging, D., Siapera, E. (2018) 'Special issue on online misogyny', *Feminist Media Studies*, vol. 18(4), pp. 515-524.

6 See https://saferinternet.org.uk/blog/revenge-porn-pandemic-rise-in-reports-shows-no-sign-of-slowing-even-as-lockdown-eases.

7 Ging, D., Siapera, E. (2018) 'Special issue on online misogyny', *Feminist Media Studies*, vol. 18(4), pp. 515-524.

8 Ging, D., Siapera, E. (2019) *Gender Hate Online: Understanding the New Anti-Feminism*, Dublin, Palgrave Macmillan.

Individual harms deriving from gender-based infringement of privacy may range from reputational impact, economic harm due to the loss of professional reputation, reduction in job or educational opportunities, discrimination in employment, stigmatisation, social isolation, restrictions in the freedom of movement, mental health issues and even suicide, physical danger, arrest, imprisonment, or execution in some jurisdictions[9].

2. Article 8 of the European Convention for the Protection of Human Rights and Fundamental Freedoms

Article 8 of the European Convention for the Protection of Human Rights and Fundamental Freedoms (henceforth ECHR) proclaims the "right to respect for private and family life"[10].

Article 8 of ECHR imposes respect even between private parties. For this reason, the Court will evaluate the eventual State's adoption of measures to assure the effective protection of those. Private life is a "broad concept", "incapable of exhaustive definition", which might "embrace multiple aspects of the persons' physical and social identity"[11].

The protection of personal data[12] is an essential requirement for the enjoyment of the right to privacy and other fundamental rights and freedoms.

Furthermore, "private life" encompasses the physical and psychological integrity of the person as well. The State' positive obligation includes a duty to enforce in practice an adequate legal framework affording protection against acts of violence by private individuals, whether the necessary severity threshold's outcome is positive. For example, an Internet user and those who provide access to a website. In other words, there is a positive obligation on the State to

9 As reported by Special Rapporteur on the Right to Privacy, 2019. Annex 2: The Human Right to Privacy: a Gender Perspective. Geneva: Human Rights Council, (A/HRC/40/63).

10 Everyone has the right to respect for his private and family life, his home and his correspondence. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

11 See Niemetz v. Germany, Application no. 13710/88, Judgment, Strasbourg, 16 December 1992; Pretty v. The United Kingdom, Application no. 2346/02, Judgement, Strasbourg, 29 April 2002; Peck v. The United Kingdom, Application no. 44647/98, Judgment, Strasbourg, 28 January 2003; S. and Marper v. The United Kingdom, Application nos. 30562/04 and 30566/04, Judgment, Strasbourg, 4 December 2008.

12 Means any information relating to an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

provide an effective deterrent against serious and grave acts at the detriment of one's personal data. Particularly, the State is responsible if it fails to enforce the necessary measures to protect victims when the risks of violence are known[13].

Article 8 ECHR includes the protection of personal data, the physical and psychological aspects of one's life, as well as one's sexual life, as broadly interpreted and applied in the "living" jurisprudence of Strasbourg. The protection of personal data is of essential importance to an individual's enjoyment of the right to respect for private life. Furthermore, the principle is that Article 8 ECHR affords protection to personal information which individuals can legitimately expect should not be published without their consent. In other words, the non-consensual dissemination of intimate images is clearly a matter of "private life", therefore calling for an application of the guarantees of Article 8 ECHR[14].

One of the most recent cases of revenge porn in Italy regards a young kindergarten teacher in Turin. Her boyfriend spread, via a WhatsApp chat between friends, a video of them in an intimate context. The chat also included the father of a little girl who attended the teacher's kindergarten, and the video was then sent to other parents and the headmistress.

As a result of this, the teacher was judged to be unprofessional. As she denounced in Court, she was immediately/later summoned by the headmistress who forced her to resign instead of defending her.

During the proceedings, several people were accused of defamation and private violence. The ex-boyfriend, the main accused of spreading the video, defends himself by saying that it was the girl who made a mistake sending the video to him because "certain things a teacher should not do".

The issue of revenge porn inevitably affects the fundamental right to privacy, as well as the victim's honor, reputation and freedom of sexuality. In Italy, the Data Protection Authority also plays an important role, whose activity complements that of judicial authorities. To make the role of the Data Protection Authority even more important and effective, Law Decree No. 139/2021, converted by Law No. 205/2021, introduced the new Article 144 bis of the Privacy Code. The Article states that "anyone, including minors over the age of 14, who has well-founded reason to believe that images or videos with sexually explicit content concerning him or her, intended to remain private, may be sent, delivered, transferred, published or disseminated without his or her consent in violation of Article 612-ter of the Criminal Code, may contact the Data Protection Authority by means of a report or complaint".

13 Stringhi, E. (2020) *Revenge Porn On Online Platforms: Legal Interpretations And Approaches To Combat Non-Consensual Intimate Image Distribution.* Privacy e innovazione [Master Law Thesis], vol. 6, Università degli Studi di Milano.

14 Ibid.

The new Article 144bis can be considered a very important step for the protection of victims – even potential victims – of revenge porn.

As a matter of fact, it is sufficient to consider the data of a survey carried out by Women for Security, a community of Italian female cybersecurity professionals. From this research – mainly composed of women's testimonies – it emerged that about 50% of victims of revenge porn decide not to denounce it. The reasons are similar to those concerning victims of other sex crimes: shame, the fear of being judged as those who brought it on themselves, anxiety about the judgments of family and friends, and the burden of criminal proceedings.

These considerations allow us to understand how important the work of the Data Protection Authority can be. In fact, the Data Protection Authority has faster procedures than judicial ones. Moreover, cooperation with the public ministry makes it possible to maintain a kind of contact between the proceedings for the protection of privacy and those for the repression of criminal conduct[15].

3. Conclusion

"Non-consensual dissemination of intimate images", as an ulterior conceptualisation of the phenomenon, constitutes an intuitive, encompassing, precise and neutral legal notion. It focuses on the invasion of sexual privacy and on the lack of consent of the subject whose private intimate images were created or obtained, after distribution, without leaving space for gender-biased assumptions, nor victim blaming. On this connection, the argument of fairness to offended subjects justifies the use of the terms "image-based sexual abuse" or "non-consensual dissemination of intimate images", within the theorised conceptual framework of "fair labelling" in criminal law[16].

Persistent online attacks disproportionately target women and frequently include detailed fantasies of rape as well as reputation-ruining lies and sexually explicit photographs. And if dealing with a single attacker's "revenge porn" were not enough, harassing posts that make their way onto social media sites often feed onto one another, turning lone instigators into cyber-mobs. Cyberharassment is a matter of civil rights law, Citron contends, and legal precedents as well as social norms of decency and civility must be leveraged to stop it[17].

---

15 See https://www.altalex.com/documents/news/2022/07/26/revenge-porn-primi-provvedimenti-tutela-potenziali-vittime.

16 Williams, G. (1983) 'Convictions and Fair Labelling', *Cambridge Law Journal,* vol. 42(1), pp. 85-95; Chalmers, J., Leverick, F. (2008) 'Fair Labelling in Criminal Law', *The Modern Law Review,* vol. 71(2), pp. 217-246.

17 Citron, D.K. (2014) *Hate crimes in cyberspace*. Harvard University Press. Available at: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=vfB7fyUAAAAJ&citation_for_view=vfB7fyUAAAAJ:9ZlFYXVOiuMC.

Chapter VI
On the Relationship Between Competition Law and Privacy: Can we Achieve Nexus Between Competition Law and Privacy?

by Arletta Gorecka*

INDEX: 1. Confusing relationship between competition law and data privacy law: background. – 2. How privacy could be relevant for competition law? – 3. Achieving a nexus between competition law and privacy.

1. Confusing relationship between competition law and data privacy law: background

Traditionally, competition law and privacy have been seen as separate. The influence of GAFAM companies has blurred the divide between privacy and competition law regulation. The current EU-level approach to privacy infringements, as per the case of Asnef-Equifax, indicates that: "any possible issues relating to the sensitivity of personal data are not, as such, a matter for competition law, they may be resolved on the basis of the relevant provisions governing data protection"[1]. This reasoning is correct on the basis that the mere breach of data protection provisions should not be seen as a competitive matter.

The advent of big data fueled innovation, leading to the emergence of new products, services, and business models. With the influence of the GAFAM companies, the divide between competition law and privacy has been blurred. Essentially, both legal orders have been subject to challenges due to the digital transformation, and the economic growth of the firms, with a question remaining on defining the level of interaction between competition law and privacy.

---

* PhD in Competition Law (University of Strathclyde), specialising in the digital economy and privacy. With an academic career, Arletta has imparted knowledge on various legal subjects (including Scots law subjects, Media Law, Internet Law and Telecommunication Law) at multiple universities across Scotland. Currently, Arletta Gorecka serves as a Module Coordinator and Law Tutor at Glasgow International College. Additionally, Arletta Gorecka is an active member of the Competition Law Sub-Committee at the Law Society of Scotland and has authored several notable publications in the field of competition law.

1 Case C-235/08 Asnef-Equifax v. Asociación de Usuarios de Servicios Bancarios ECR I-11125. [2006]; para 177.

Interaction between competition law and privacy is new. Lack of control over personal data appears to be the "core theme" for consumers[2], and the market functioning. Fundamentally, competition law is a tool for managing the private market: that is, the exchange of goods and services for value between individual particulars in the market. Indeed, the indirect benefits of competitive markets for social stability and democracy all motivate the strong commitment of the EU to the competition law. Data protection and competition policies share foundational concerns and similar remedial approaches: how to mitigate unfairness by introducing and imposing obligations on those with information or market power. The goal is to prevent power imbalances between individuals and powerful companies. However, data protection legislation offers only a partial response to the exploitation of personal data. Data protection law does not recognise the long-term harms to the platforms' users, including the special responsibility that could ensure the position of power of some digital platforms.

Yet, for competition lawyers, the problem relates to what extent competition law could resolve the issues of privacy. I argue that for competition law, it is of utmost importance that the competition problem is interlinked to the privacy dilemma. Firstly, the economic characteristics of online digital platforms, and their monopolistic features, allow for unprecedented data collection due to the lack of consumer choice. Secondly, such broad access to personal data could further entrench the economic power of digital firms via increased barriers to entry, and a possibility to use informational asymmetry between consumers and them, as well as behavioural manipulation digital service providers use. Hence, this role of personal data in the digital gatekeepers' business models also leads to privacy concerns being interlinked with competition policy problems.

2. How privacy could be relevant for competition law?

Digital markets have unquestionably linked competition law with data privacy law concerns. Lack of control over personal data appears to be the "core theme" for consumers, and the market functioning. Fundamentally, competition law is a tool for managing the private market: that is, the exchange of goods and services for value between individual particulars in the market. Indeed, the indirect benefits of competitive markets for social stability and democracy all motivate the strong commitment of the EU to the competition law. Data protection and competition policies share foundational concerns to prevent power imbalances between individuals and powerful companies.

Privacy protection does not have a recognisable definition, which encompasses an importance of personal privacy to their political affiliation, social

2 CMA, The Commercial Use of Consumer Data – Report on the CMA's Call for Information (2015).

life, or dignity. The concept of privacy might be defined both in a broad and narrow sense. Unquestionably, any breaches of privacy affect a great number of peoples and could potentially compromise the process of democracy.

There are two most articulated, yet opposing, theories on the interface between competition law and privacy. Nonetheless, the theories are new as the intersection between competition law and privacy is quite new. Essentially, this originates with the reasoning behind understanding what privacy means. Its definition has proven to be a divisive and slippery concept. Warren and Brandeis influenced the reasoning and defined the conception of privacy as "the right to be let alone"[3]. However, the modern collection of privacy rights and interest left little of broadly accepted definition of privacy. The modern understanding of privacy provides its ambitious nature. Scholars acknowledged the concept of privacy as "chameleon-like", "in disarray" and "vague and evanescent"[4].

The literature recognises two opposing views on the intersection between competition law and privacy. The first theory considers data protection law as beyond considerations of competition law[5]. Arguably, the separatist view originates from the Asnef-Equifax case, where the court rejected the intersection between competition law and privacy. Essentially, separatist theory views competition law and privacy as "complementary [in] nature", but not overlapping[6]. The central argument remains that incorporation of privacy concerns in competition law assessment would create confusion, especially in application to consumer welfare standard.

On the contrary, the integrationist approach accepts incorporation of privacy into the longstanding competition law framework[7]. Based on that, consumer welfare could be improved by consideration of both price and non-price factors[8]. Essentially, when there is evidence that companies compete to offer privacy to their consumers[9], the integrationist approach considers if privacy-based competition might be impacted. To picture this effectively, one could

3 Warren S.D., Brandeis, L.D. (1890) 'The Right to Privacy' *Harvard Law Review*, vol. 4(5).

4 Solove, D.J. (2006) 'A Taxonomy of Privacy', *154 U. PA. L. REV.*, pp. 479-80. Available at: https://scholarship.law.upenn.edu/penn_law_review/vol154/iss3/1 (noting these many observations of the difficulty in defining privacy); Fairfield, J.A.T., Engel, C. (2015) 'Privacy as a Public Good', *Duke Law*, vol. 65(3). (Privacy theorists differ famously and widely on the proper conception of privacy).

5 See Cooper, J.C. (2013) 'Privacy and Antitrust: Underpants Gnomes, the First Amendment, and Subjectivity', *George Mason Law Review*, pp. 1129-1146.

6 Ohlhausen, M.K., Okuliar, A.P. (2015) 'Competition, Consumer Protection, and the Right [Approach] to Privacy', *Antitrust Law Journal*, pp. 138-143. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2561563.

7 Douglas, E.M. (2021) 'The New Antitrust/Data Privacy Law Interface', *The Yale Law Journal Forum*, p. 647.

8 See for example, National Soc'y of Prof. Engineers v. United States, 435 U.S. 679 (1978).

9 Pasquale, F. (2013) 'Privacy, Antitrust, and Power', *George Mason Law Review*, vol. 20(4), pp. 1009-1024.

consider a hypothetical merger of two internet-based companies. If before this merger, these companies compete to offer different levels of privacy, the assessment might consider if their merger substantially reduces the privacy options available to consumers post-merger. Integrationist approach would assess if privacy-as-quality's reduction leads to reduction of competition. However, if there was no privacy-as-quality competition between merging parties, then integrationist approach would consider any privacy related concerns to be beyond the competition law assessment[10].

The German Facebook case[11] remains the example of a competition authority diminishing the boundaries between competition and data protection law. Germany has been the most active in integrating privacy into competition law. Its unique perspective has provided the most innovative approach to acknowledge privacy concerns into the exploitative competition law. At the EU level, AG Rantos opined that an incidental consideration of GDPR could be admissible for competition law assessment[12]. This could only be achieved if the GDPR is introduced into a wider scope of the legal and economic context surrounding the conduct[13]. Such argument has also been accepted in the Facebook case, where the CJEU adopted AG Rantos' line of reasoning. The CJEU emphasises that the breach of an area of law can play a role in assessing a possible violation of competition law, as recognised in cases of AstraZeneca and Allianz Hungaria[14]. Hence, an incidental consideration of GDPR in cases of possible competition law infringement should not be seen as unexpected. Correspondingly, the CJEU concurred that while compliance with the GDPR has not automatically ruled out the finding of an abuse, it can be taken into consideration as part of the comprehensive assessment[15]. In this context, it may serve as a crucial indicator in determining if the conduct employs strategies in accordance with merit-based competition.

Acquisition of the data on its own is not abusive and could promote competition. However, on a practical level, acquisition of personal data could introduce competitive constraints. Firstly, the economic characteristics of online

10 See Statement of Federal Trade Commission Concerning Google/DoubleClick, FTC File No. 071-0170, at 2 (F.T.C. Dec. 20, 2007).

11 Case B6-22/16 Facebook, Exploitative business terms pursuant to Section 19(1) GWB for inadequate data processing. Available at: https://www.bundeskartellamt.de/SharedDocs/Entscheidung/EN/Fallberichte/Missbrauchsaufsicht/2019/B6-22-16.html?nn=3600108.

12 Case C-252/21 Meta Platforms and Others (Conditions Générales d'Utilisation d'un Réseau Social) Opinion of AG Eantos, para 24.

13 Ibid, para 23.

14 Case C-252/21 Meta Platforms and Others (Conditions Générales d'Utilisation d'un Réseau Social), para 110; C-457/10 P AstraZeneca/Commission ECLI:EU:C:2012:770; C-32/11 Allianz Hungária Biztosító and Others ECLI:EU:C:2013:160.

15 Case C-252/21 Meta Platforms and Others (Conditions Générales d'Utilisation d'un Réseau Social), paras 62 and 110.

digital platforms, and their monopolistic features, allow for unprecedented data collection due to the lack of consumer choice. Secondly, such broad access to personal data could further entrench the economic power of digital firms via increased barriers to entry, and a possibility to use informational asymmetry between consumers and them, as well as behavioural manipulation digital service providers use. Hence, this role of personal data in the digital gatekeepers' business models also leads to privacy concerns being interlinked with competition policy problems. The crucial element of an antitrust infringement is some form of misconduct that violates competition. In other words, it is the conduct that not merely involves data mistreatment, and breach of users' privacy, but the conduct must have a negative impact on the competition. The Internet's cornerstone is personal data, which is a necessary component of digital platforms' business models, such as Facebook and Google[16]. The digital platforms acquire and treat data as a necessary component to improve the quality of their services, aiming at enhancing their attractiveness to existing and potential customers. To this effect, considering the platform's business model is necessary. The data protection consent-based framework benefits online platforms, creating the incentives for functional consent to engage in data processing activities. By including confusing and deceptive terms in privacy policies as well as engaging in unfair commercial practices, online platforms contribute to mandating the status quo in the online markets. They keep consumers confused or uninformed as to the privacy-related implications of accessing and using their online services and products, deepening informational asymmetry which affects digital markets. Accordingly, under constraints of privacy, the collection of data continues to deepen platforms' market power. There could be two ways in which privacy could intersect or influence competitive theory of harm: increased and decreased privacy protection strategy by Big Tech firms. The element of increased privacy protection deals with requiring users to consent to data collection (Apple). This reasoning has been noted in the BKA's Facebook case. Consumers might be well exploited by being offered "zero price" in terms of monetary transactions. Such zero price-reasoning could be arbitrary and underline market failure in the acquisition of private personal data. Current privacy regulations ignore this market failure as they are based on consumers "rights" and ignored that something might be wrong with this market, which is framed by these rights. The behavioural problems, a combination of personal data and the market power could endanger the excessiveness of personal data collection, which is a problem for exploitative abuse of market power as well as for the informational self-determination and protection of users' privacy. In

16 Reding, V. (2012). The EU Data Protection Reform 2012: Making Europe the Standard Setter for Modern Data Protection Rules in the Digital Age. Innovation Conference Digital, Life, Design. (Munich, 2012, SPEECH/12/26). Available at: https://ec.europa.eu/commission/presscorner/detail/en/SPEECH_12_26.

this respect, due to practices of large digital undertakings, we have an unsolved privacy problem, which is a challenge for data protection law.

3. Achieving a nexus between competition law and privacy

In the context of the debate on defining the nexus between competition law and privacy, only one question remains unanswered: to what extent could privacy-related harms be seen as directly influencing competition law assessment?

The interactions between competition law and privacy are nascent and complex. Despite the common denouncement that the relationship between competition law and privacy is complementary, the relationship between competition and data privacy is much more multi-faced and nuanced. The Author acknowledges that the intersection between competition law and privacy is new. It is important to note that competition law and data protection law are two different and separate legal orders. However, data protection and competition policy share foundational concerns and similar remedial approaches: how to mitigate unfairness by introducing and imposing obligations on those with information or market power and giving these rights to those with fewer powers. Essentially, the goal is to prevent power imbalances between individuals and strong companies.

Digital-consumers-oriented markets are characterised by weak competition, and widespread confusion about the privacy-related consequences of the T&Cs offered by digital platforms and service providers. Consumers are exploited by the digital platforms as they are unable to act upon the offered T&Cs due to informational asymmetries and their bounded rationality. The overall strategy of the large digital companies analysed in this thesis is to acquire and process a high volume of data through lawful and unlawful means. I argue that the competition law authorities should focus on establishing the scenarios where the processing of data gives a rise to anticompetitive effects. To this effect, they need to determine: i) the kind of data collected and proceeded that likely contribute to the strengthening of a dominant position, and ii) the fact that the data could have been collected and processed unlawfully.

Firstly, not all types of data could attain to strengthen a dominant position of digital undertakings. To determine what kind of acquired data could contribute to the attainment of a dominant position, it is important to focus on specific markets. For instance, Google is dominant in the search and search advertising markets, as well as Facebook is dominant in social networking and social network advertising markets. If we focus on Google's case, Google needs to render more detailed and relevant search results for users and provide targeted advertisements for the advertising side. Google's strategy has allowed for a high personalisation of the offered search services. In turn, personalisation

introduces the consolidation of users' habits[17], leading to creating lock-in effects and informational asymmetry.

The effect of such misleading practices is the preservation of the gatekeeper's dominance. However, how should competition law deal with this balancing practice? I recommend the use of a cautious approach. The problems of competition law and privacy-related harms are interlinked, and competition law should only remediate such concerns that directly harm competition in a relevant market. This approach acknowledges competition law as being limited to competitive issues but postulates that privacy protection could form a competition dimension[18]. The requirement to consider privacy as a competition dimension is important in consideration of services offered at zero price in exchange for personal data. However, this approach could have been further argued to show methodological concerns, since a circle of data procession by an undertaking would indicate a possibility of market failure and corresponding privacy harm. A particular problem is that gatekeepers would always try to get higher users' attention and their consent. In turn, this could lead to degrading of the offered service or product's quality which takes a form of poorer privacy protection. The combination of data sets could allow for a deeper consumer profiling, but it is not clear if consumers could assess any additional risks. Based on that debate, the Author proposes that the following. Firstly, competition law and data privacy are overlapping in complex and multifaceted ways particularly in the digital economy. However, this does not mean that competition law should be extended to cover every data protection problem. However, competition law should intervene in conduct involving privacy-related theories of harm only if privacy-related harm relates to the market failure, not to the privacy rights itself.

Competition law is concerned with market power that might negatively impact consumer welfare. The current EU-level approach to privacy infringements emphasises that data protection law is outside the scope of competition law. The Author agrees with this reasoning, on the basis that the mere breach of data protection provisions should not be seen as a competitive matter. While privacy standards are relevant to competition analysis as a qualitative parameter and can play a part in competition assessment, it is important to keep competition law and data protection distinct. This should not constitute an expansion of the competition legal order to encompass other areas of law, such as data protection, for the simple reason that competition law relates to tacking harms caused by market failures and seeks to remedy competitive harm. Competition law should continue to support the prevailing approach and assess decisions

17 Paal, B.P. (2016) 'Internet Search Engines and Antitrust law', *Journal of Intellectual Property Law & Practice*, vol. 11(4), pp. 298-301.

18 Stucke, M., Grunes, A. (2016) Introduction: Big Data and Competition Policy, Oxford University Press.

involving personal data through the spectrum of protecting a competitive equilibrium in hypothetical markets.

However, with the respect to the economic power of the gatekeepers, the competition problem cannot be analysed independently without considering the behavioural and informational problems relating to the acquisition of personal data.

Reduction in privacy may not necessarily amount to a competitive issue; any reduction of privacy equally may not immediately breach data privacy law, if the data processes comply with data protection law. Competition law has a role to play in promoting users' privacy. Even if we assume that privacy itself might not be easily reduced to a commodity on the markets, privacy protectionism will be recognised as an element of a product's quality. To the extent that individuals value privacy, protection competition enables consumers to choose better privacy. By promoting competition in the markets, competition enforcement might ensure consumers can make real choices among products and services.

On the fundamental level, competition and data protection laws achieve different sets of modalities. Competition law aims to ensure undistorted competition within the internal market. Competition is conceived as the best means to ensure allocation of resources and increase consumer welfare. Privacy distortion effects could arguably both distort fairness and efficiency-based theories. Yet, we are only at the beginnings to categorise this approach. If it holds true, then it will be incredibly convenient tool, creating a cohesive legal system. The system could pursue its respective enforcement goals.

Arguably, competition law could act as an effective tool for protecting privacy. To the extent consumers should remain informed and choose products or services offering robust privacy protections, competition law is integral to protecting consumers' choices. By recognising the aggregation of personal data as a potential source of market power, competition law enforcement might provide recourse where companies use their market power to inflict harm degrading privacy. However, while privacy standards are relevant to competition analysis as a qualitative parameter, it is important to keep competition law and data protection distinct. Essentially, maintaining the analytical independence of legal orders could contribute to the achievement of predictability and ensures respect for legality.

Chapter VII
Assessing Risks Involved in the Use of AI Systems: Current and Future Approaches

by Pietro Boccaccini* and Taís Fernanda Blauth**

INDEX: 1. The risks of Artificial Intelligence. – 2. Assessing AI Risks through an ethical perspective. – 3. Operationalising AI risks assessments. – 4. Future approaches to assessing AI Risks. – 4.1. Fundamental rights impact assessment (FRIA) and data protection impact assessment (DPIA). – 4.2. Risk management system. – 4.3. Transparency obligations as a guide to carry out a DPIA. – 4.4. Technical documentation. – 4.5. EU declaration of conformity. – 4.6. Incidents reporting. – 5. Conclusion.

1. Unpacking the risks of Artificial Intelligence

Artificial Intelligence (AI) systems have been implemented in most, if not all, sectors of society. Its integration into diverse sectors – from healthcare to finance, and national security – demonstrates unparalleled utility. However, the versatility of AI also opens avenues for misuse and abuse, raising ethical, security, and governance challenges. Among the malicious uses of AI are the forgery of videos and images, implementation of AI in weapons systems and advanced social engineering techniques with the purpose of accessing personal data. In addition, AI systems can also be abused by ill-intended actors. Such actions include the manipulation of the stock market leading to *flash crashes* and membership inference attacks to uncover protected data. Moreover, the use of AI possibly also implies the risk of bias in predictive analyses, discrimination, and large-scale surveillance. Should personal data (or special categories of data) also be processed by the system, the risk level significantly increases.

Artificial Intelligence is one of the so-called "dual-use technologies", given it can be a tool or a weapon[1]. This dual-use nature is not unique to AI but

---

* Lawyer with 15 years of professional experience specializing in privacy, data protection, cybersecurity and artificial intelligence (AI). Since 2021, he has served as the coordinator of Deloitte Legal's Data Protection Team in Italy. Pietro's expertise extends to acting as a Data Protection Officer (DPO) for groups operating in diverse sectors. In addition to his professional commitments, he actively contributes to the legal community as an expert for the European Data Protection Board (EDPB), a lecturer in Legal Informatics at the University of Milan and a residential affiliate of the Information Society Law Center (ISLC). He is the author of several scientific publications on tech matters.

is emblematic of technological advancements throughout history, such as the internet and nuclear energy. If one considers technology more broadly, it is also the case for other tools that are usually not classified as a "technology", such as hammers. A hammer is a handy invention that enables people to fix furniture, build houses or hang a painting on the wall. However, in the hands if an ill-intended person, it can harm or even kill. Given this scenario, one might wonder why society is suddenly so concerned with the risks posed by AI, when there are so many other technologies that can also be considered dangerous. Why was the same attention, or even hype, not given to other technological advancements? Why does AI, in particular, incite significant concern and scrutiny?

AI, for the better or for the worse, has its own particularities. It can reach a much larger scale and reach than other technologies. With the use of techniques such as Machine Learning (ML), the initial capabilities of a system can be greatly expanded, and an algorithm can even respond to new data in unpredictable ways. With the increased use of such techniques in everyday systems, in both governmental and private institutions, citizens are directly affected by errors or intentional malicious uses of AI. To understand and assess the possible risks, it is useful to consider two main categories[2]:

1. Malicious use of AI: referring to "the use of AI to enhance, augment, or enable acts committed by individuals or organizations. This includes practices not necessarily considered crimes by specific legislation, but that still compromise the safety and security of individuals, organizations, and public institutions"[3]. Some of the malicious uses of AI include the creation of deepfakes, autonomous weaponry, and sophisticated social engineering attacks aimed at personal data exploitation. Such uses not only threaten individual privacy and security but also challenge societal norms and international peace. Additional risks can be seen in the figure below:

** Taís Fernanda Blauth received a degree in Law, an MA degree in international politics, and she is currently a PhD researcher in Artificial Intelligence and International Relations at the University of Groningen/Campus Fryslân, The Netherlands. She is a member of the Data Research Centre and a fellow at the Humboldt Institute for Internet and Society, in Germany, and the Information Society Law Center (ISLC), in Italy.

1 Smith, B., Browne, C.A. (2019) *Tools and Weapons: The Promise and the Peril of the Digital Age*. Hodder & Stoughton.

2 Blauth, T.F., Gstrein, O.J., Zwitter, A. (2022) 'Artificial Intelligence Crime: An Overview of Malicious Use and Abuse of AI', *IEEE Access*, vol. 10, p. 77110. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4190323.

3 Ibid.

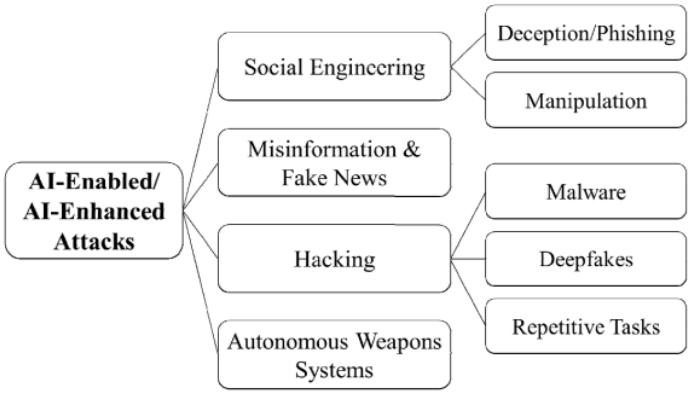

Figure 1. Malicious Use of AI[4]

2. Malicious abuse of AI: referring to "the exploitation of AI with bad intentions, as well as attacks on AI systems themselves"[5]. Examples, also shown in the figure below, include market manipulation through algorithmic trading leading to financial instability, and membership inference attacks that compromise data privacy. This dimension also encompasses attacks on the AI systems themselves, undermining their integrity and reliability.

4 Ibid (p. 77118).expanding existing vulnerabilities, and introducing new threats. This article reviews the relevant literature, reports, and representative incidents which allows to construct a typology of the malicious use and abuse of systems with AI capabilities. The main objective is to clarify the types of activities and corresponding risks. Our starting point is to identify the vulnerabilities of AI models and outline how malicious actors can abuse them. Subsequently, we explore AI-enabled and AI-enhanced attacks. While we present a comprehensive overview, we do not aim for a conclusive and exhaustive classification. Rather, we provide an overview of the risks of enhanced AI application, that contributes to the growing body of knowledge on the issue. Specifically, we suggest four types of malicious abuse of AI (integrity attacks, unintended AI outcomes, algorithmic trading, membership inference attacks).

5 Ibid (p. 77112).

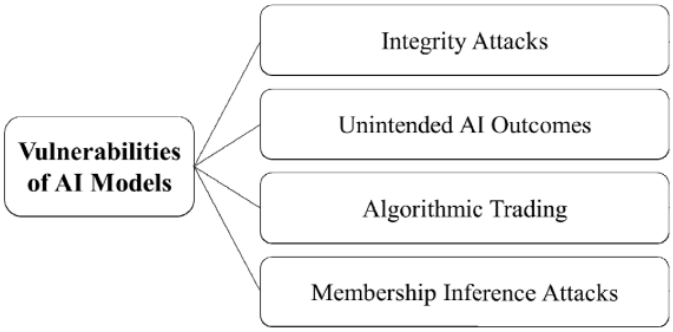

Figure 2. Malicious Abuse of AI[6]

As technology develops and techniques are refined, new challenges and risks can come to surface. For this reason, attempts to classify the risks are limited, but still helpful to understand the challenges at hand and consider policy strategies to deal with them.

Given this risk scenario, governance mechanisms and legislation are paramount to addressing the challenges. Even though provisions regulating data and technology exist in the European context, those are not sufficient to address the challenges raised by AI. For this reason, the European Commission started making efforts to regulate such systems. From broad and more general ethical guidelines to an advanced version of the AI Act (AIA)[7] proposal, many aspects and understandings have evolved. Together with these advances, a sense of urgency could be noticed, which might indicate a desire to establish an extraterritorial application of the AIA, following the example of the GDPR.

When regulating technology, three main approaches are considered.

6 Ibid (p. 77118).expanding existing vulnerabilities, and introducing new threats. This article reviews the relevant literature, reports, and representative incidents which allows to construct a typology of the malicious use and abuse of systems with AI capabilities. The main objective is to clarify the types of activities and corresponding risks. Our starting point is to identify the vulnerabilities of AI models and outline how malicious actors can abuse them. Subsequently, we explore AI-enabled and AI-enhanced attacks. While we present a comprehensive overview, we do not aim for a conclusive and exhaustive classification. Rather, we provide an overview of the risks of enhanced AI application, that contributes to the growing body of knowledge on the issue. Specifically, we suggest four types of malicious abuse of AI (integrity attacks, unintended AI outcomes, algorithmic trading, membership inference attacks

7 On 21st April 2021 the European Commission published a proposal to regulate artificial intelligence in the European Union. On 6 December 2022 the Council of the EU adopted its common position on the AI Act. The proposal of the AI Act will become law once both the Council (representing the 27 EU Member States) and the European Parliament agree on a common version of the text. The final draft of the proposal can be found in the following link: https://artificialintelligenceact.eu/the-act/.

1. The sectoral approach, traditionally followed in the United States, focuses on establishing rules based on specific sectors. For instance, creating privacy and data protection legislation to protect health data.
2. The omnibus approach, also known as the principle-based approach. More common in European regulations, it helped to shape the GDPR framework.
3. The risk-based approach, which creates different provisions depending on the risk level of technologies.

In the AIA proposal, it is possible to identify a principle-based approach but, more evidently, the Regulation is framed around different levels of risk.

2. Assessing AI Risks through an ethical perspective

The rapid advancement of artificial intelligence (AI) technologies poses unique ethical challenges and risks, necessitating a nuanced approach to governance that extends beyond conventional regulatory frameworks. The creation or adaptation of laws is a lengthy process, which is not always able to keep up with the speed of technological advances. In this dynamic landscape, ethical principles, and best practices emerge as pivotal tools for guiding responsible AI development and application, particularly during periods of legislative inertia or societal upheaval. In this scenario, ethical principles and good practices are useful and crucial instruments for the development and responsible use of technologies. They are especially relevant in times of political, economic, or social instability. Historical precedents, such as the digital solutions deployed during the COVID-19 pandemic, underscore the significance of ethical considerations in the absence of specific legal guidelines. These solutions, designed to manage the pandemic's challenges through contact tracing, population movement monitoring, and public health information dissemination, sparked considerable debate over privacy and transparency. Such instances highlight the crucial role of ethical principles in safeguarding individual rights and public trust, especially when regulatory mechanisms lag behind technological innovations[8]. Therefore, ethical principles are relevant, especially in the absence of specific legislation. Over the past decade, these principles have been central to debates about risk reduction related to emerging technologies. For example, companies, research institutes, civil society organizations, and government institutions have published numerous documents containing guidelines for "ethical artificial intelligence". However, due to the large number of publications and principles, the most relevant elements are not always clear. To identify if there is agreement,

8 Blauth, T.F., Gstrein, O.J. (2021) 'Data-Driven Measures to Mitigate the Impact of COVID-19 in South America: How Do Regional Programmes Compare to Best Practice?', *International Data Privacy Law.* Available at: https://doi.org/10.1093/idpl/ipab002.

at a global level, regarding ethical requirements and good practices, Jobin et al. evaluated 84 of these documents[9].

The results of the analysis demonstrate that there is a convergence around 5 principles:

1. Transparency: To ensure that AI algorithms are reliable, it is necessary that the capabilities and purpose of AI systems communicated openly and decisions – as far as possible – explainable to those affected directly and indirectly.
2. Fairness and equity: the development and use of AI must be fair, which includes preventing and mitigating unwanted biases that arise from algorithms, enabling the right to challenge decisions of AI systems, as well as providing fair access to technology and its benefits.
3. Non-maleficence: related to the need to ensure that systems are secure. For this, it is necessary to avoid some risks or potential damages, such as violation of privacy, physical damages, damages arising from discriminatory algorithms, and misuse of AI in cybercrime, among others.
4. Responsibility and accountability: involve the importance of assigning legal responsibility in case of problems related to AI systems. It is ideal for this responsibility to be clarified in advance (for example, in contracts). If this is not possible, it should be possible to identify it as a way to remedy any damage.
5. Privacy: seen as a value and a right to be protected when talking about the development and ethical use of AI. It is usually linked to the right to data protection and data security.

An initial analysis may indicate that the large number of documents reinforcing ethical guidelines is a positive aspect. However, what is perceived is that, in many cases, the proliferation of ethical guidelines has served as a façade for ethics washing[10], rather than a legitimate intention to create governance instruments for emerging technologies. In practice, many companies that promote initiatives such as AI for good[11], end up using these systems for their own benefit, such as "surveillance capitalism"[12].

This problem was highlighted in 2018 when thousands of Google employees signed a letter protesting the company's involvement with the Pentagon in the

9 Jobin, A., Ienca, M., Vayena, E. (2019) 'The global landscape of AI ethics guidelines', *Nature Machine Intelligence*, vol. 1, pp. 389-399.

10 Johnson, K. (2019) 'How AI Companies Can Avoid Ethics Washing', *VentureBeat*. Available at: https://venturebeat.com/2019/07/17/how-ai-companies-can-avoid-ethics-washing.

11 See https://aiforgood.itu.int.

12 Zuboff, S. (2019) *The Age of Surveillance Capitalism*, Profile Books.

Maven project[13]. This project was an initiative of the United States Department of Defense, in which Google would contribute with an improvement in the object recognition process in drones, using AI techniques. One of the employees' criticisms is related to the company's self-image, which used the motto "don't be bad", but which contributed to the development of technologies that can be used in weapons and wars. After the negative repercussions, the technology giant decided not to renew the contract with the Pentagon. This event shows that, as much as companies and governments adopt a list of ethical principles to guide the use of disruptive technologies, such principles and slogans are not always enough.

The challenges in aligning AI development with ethical standards underscore the necessity for a comprehensive legislative framework that incorporates ethical principles into legal mandates. Initiatives in this sense can already be seen in some countries. Within the European Union, for example, the regulation of AI systems has been one of the priorities in recent years and the legislative proposal is at an advanced stage, with the AIA.

It shall be reminded that the European Commission's objectives in preparing the AIA are mainly the following: ensuring that AI systems used in the EU are safe and respect existing law on fundamental rights and European values, ensuring legal certainty in view of facilitating investment and innovation in AI, enhancing governance and effective enforcement of existing law on fundamental rights and safety, and facilitating the development of a single market for lawful, safe and trustworthy AI applications, while preventing market fragmentation.

3. Operationalising AI risks assessments

As a general remark, it could be argued that the use of AI: (i) from one hand, has been facilitated in last years by the absence of specific rules, that generally determine longer processes to start and run projects; (ii) from the other side, might have been hindered by the lack of legal provisions, that help in approaching the risks following a given methodology.

In any event, also before the EU Institutions reached an agreement on a specific regulation on AI, additional rules should have been considered and remain fundamental. In particular, the EU Regulation 679/2016, the General Data Protection Regulation ("GDPR") introduced several principles and rules that can be applied in an AI context. These provisions must be respected in all the cases in which personal data are processed by the "machine".

13 Shane, S., Wakabayashi, D. (2018) 'The Business of War": Google Employees Protest Work for the Pentagon', *The New York Times*. Available at: https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html.

Under the GDPR, the likelihood and severity of the risk to the rights and freedoms of the data subject should be determined based on objective assessments, also considering the nature, scope, context, and purposes of the processing activities[14]. To enhance compliance with the GDPR, where processing operations are likely to result in a high risk to the rights and freedoms of natural persons, the controller shall carry-out a data protection impact assessment (DPIA) to evaluate the origin, nature, particularity, and severity of that risk[15].

Carefully evaluating the risks is of paramount importance considering applicable rules, given that technical and organisational measures that must be adopted pursuant to the GDPR are those necessary to ensure a level of security which is appropriate to the risk.

The DPIA methodology is an extremely flexible tool (Article 35, GDPR). The mentioned article lists certain: i) cases where the assessment is required[16]; and ii) elements that shall necessarily be included in the DPIA[17]. The provisions, however, leave open the possibility to extend the use of the assessment also in different cases and to structure the analysis using additional criteria. It could, therefore, be argued that today the DPIA to be carried out according to the GDPR is still a very effective tool, if personal data are processed.

The Article 29 Working Party (now European Data Protection Board), provided in its Guidelines[18] a concrete set of processing operations that require a DPIA due to their inherent high risk. One of the criteria to be considered is the one concerning "Innovative use or applying new technological or organisational solutions", because the use of such technology can involve novel forms of data collection and usage, possibly with a high risk to individuals' rights and freedoms: indeed, the personal and social consequences of the deployment of a new technology may be unknown. A DPIA can help the data controller to understand and to treat such risks.

The EU data protection Supervisory Authorities, in any event, are considering processing activities carried out using AI (or other innovative technologies/tools, like IoT, virtual assistants, wearable devices, wi-fi tracking, etc.) as subject to a prior DPIA[19].

---

14 Recital 76 of the GDPR.

15 Recital 84 of the GDPR.

16 Article 35, paragraph 3, of the GDPR, states when data protection impact assessment shall be required.

17 Article 35, paragraph 7, of the GDPR, states the minimum requirements of the assessment.

18 Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, adopted on 4 April 2017 and last revised and adopted on 4 October 2017 (WP 248 rev. 01).

19 The Italian Supervisory Authority (Garante) has included the processing activities carried out through AI systems in the list of those in relation to which a DPIA must be carried out (Annex 1 to the Decision 467 of 11 October 2018), if at least other criteria among those mentioned by the WP29 in the Guidelines of 2017 on DPIA is present.

Coming to the analysis through a DPIA of the potential risks to individuals' rights and freedoms connected with the use of AI, below are some operational considerations.

First, the subject required to carry out the assessment is the data controller. In practice, it may happen that a company providing certain services (e.g. interview tools to evaluate candidates' soft skills through the analysis of videos and texts) could base its activity on the use of AI: even if it acts as data processor – processing data on behalf of its clients – certain processing activities may be carried by the service provider, "owner" of the AI, also for its own purposes, as controller. For instance, training data to allow the improvement of the algorithm's function is a necessity of the service provider, not of the client, in the example at hand. Therefore, the vendor could be obliged to perform its own assessment on the processing operations carried out as controller. This subject is generally in a better position to analyse the potential risks and the related security measures, given that he has the control on the "machine", while the client using the AI needs to receive most of the information on the security measures from the service provider.

In this context, the service provider can make available to its client's part of the DPIA so that the latter can easily carry out its own assessment: this cooperation among clients-controllers and vendors-processors has become a sort of best practice lately in tech environments. Another aspect that could properly be considered in the context of a DPIA concerning an AI system is that data used by an algorithm have mainly three definitions and roles: i) input data are entered into the algorithm and used to make a prediction; ii) training data are used to generate the algorithm and train the AI; iii) feedback data are used to increase the algorithm performance with the experience. In certain cases, the same data can have these three roles. Different definitions of these categories can also be used. This classification may help in focusing the risks more precisely in the context of the assessment and, therefore, could become a possible criterion in the context of an assessment carried out considering the GDPR and related guidelines.

Personal data may be found in the following stages of the AI-based solution life cycle[20] that should also be considered while assessing risks:

1. Training: when the training stage involves personal data processing, it shall be considered, in itself, processing of personal data;
2. Validation: may include personal data processing when using data that correspond to the actual current situation of a processing activity to determine the eligibility of the experimental model;

20 GDPR compliance of processing that embed Artificial Intelligence – An introduction, Spanish Supervisory Authority, February 2020.

3. Deployment: when distributed to third parties as a component, it may be considered that there is data disclosure when the same model includes personal data or there is a way to obtain such personal data;
4. Operation: some of the operation activities of the AI-based solution may include personal data processing (e.g. inference, decision-making);
5. Removal: may include two different extensions:
   – the AI component is removed when obsolete in all processing in which it is implemented
   – a particular user decides not to use any longer the AI component; both may have effects on data suppression, as well as service portability.

Below are some of the areas where recurring impacts on fundamental rights can be found today and that should be considered – if applicable – while performing a DPIA on processing operations carried out with the use of AI, before designing and implementing the AI solution, so that adequate measures can be effectively implemented.

1. Quality of datasets is key in training AI systems: any failure in the initial data may cause incorrect outcomes and function erroneously throughout its application period, invalidating the entire system. Ensuring that the data source is trustworthy and accurate is fundamental to prevent safety issues.
2. Bias and discrimination (also within the training dataset), with far-reaching discriminatory impacts on key aspects such as gender, racial, social and other characteristics of target groups.
3. Erosion of human agency and autonomy. Misinformation campaigns in combination with elaborated recommendation engines powered by sophisticated AI can trigger addiction and opinion manipulation.
4. Fundamental right to freedom of expression and information can be severely impacted by AI systems.
5. Right to respect for private and family life may be in danger, for instance in the context of AI-powered biometric identification and facial recognition technologies collecting data that can also be highly sensitive.
6. AI and automated decision-making used in government can impact good administration, access to justice and the right to a fair trial.
7. AI may have an impact also on other EU fundamental rights, including consumer protection and the right to freedom of assembly and association.
8. Threats to several aspects of safety and security: importantly, the risks generated by AI can emerge at various phases of the product lifecycle, from the design and development phase to the deployment and post-deployment phases.

It can be difficult to describe the processing activity of AI systems, particularly when they involve complex models and data sources. However, such a description is necessary as part of a DPIA.

It is important that the DPO or other information governance professionals (or both, if possible) are involved in AI projects from the earliest stages, establishing open channels of communication between them and the project teams. This will ensure that such consultants can identify and address relevant risks early in the AI lifecycle.

4. Future approaches to assessing AI Risks

In a scenario where different types of technologies and their risks are intertwined, it becomes necessary to find ways of establishing governance convergence among the rules and approaches defined by the AIA, GDPR, Supervisory Authorities, ethical guidelines, and best practices. Otherwise, sparse regulations might have limited effect and reach.

As specifically regards the methodology to assess risks connected with the use of AI, looking at AIA it is possible to make the following remarks.

First of all, it is important to note that the European legislator has already carried out, *ex ante*, an evaluation on the risks of AI systems, establishing which are:

a. prohibited AI practices (listed in article 5)[21] and in relation to which the risk level – considering the entire EU legal framework, principles, fundamental rights etc. – is considered unacceptable and not manageable;
b. (high-risk systems (listed in Annex III), which are subject to several requirements, also in relation to the risk evaluation and management[22].

In relation to the latter high-risk systems, the following provisions of the AIA shall be considered, among other, in the perspective at hand of evaluating and managing risks.

21 E.g., AI systems that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective to or the effect of materially distorting a person's or a group of persons' behaviour, or AI systems that exploits any of the vulnerabilities of a specific group of persons due to their age, physical or mental disability, to materially distort the behavior of a person pertaining to that group.

22 E.g., remote biometric identification systems, AI systems intended to be used as safety components in the management and operation of critical infrastructures, AI intended to be used to make decisions affecting terms of the work related relationships, promotion and termination of work-related contractual relationships, AI systems intended to be used by public authorities to evaluate the eligibility of natural persons for essential public assistance benefits and services, including healthcare services, AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score.

4.1. Fundamental rights impact assessment (FRIA) and data protection impact assessment (DPIA)

The European legislator has considered the possible overlap of the DPIA with the fundamental rights impact assessment (FRIA) required for high-risk AI systems. According to paragraph 4 of article 29a of the AIA, in fact, if any of the obligations laid down in that article are already met through a DPIA, the FRIA "shall be conducted in conjunction with that data protection impact assessment". Considering this new provision, the DPIA – even if potentially satisfactory in terms of content – cannot replace the performance of a FRIA, when the latter is necessary. At the same time, however, the relevance of the DPIA is stressed by the fact the two assessments shall be made jointly, mandatorily. Therefore, in the situations where a FRIA is necessary pursuant to the AIA, the DPIA will be certainly needed; besides, this requirement is totally logical, considering the possible consequences for the rights and freedoms of natural persons connected with the use of AI systems which are already considered by the law as incorporating high risks.

Deployers of high-risk AI systems, upon the registration of such systems according to article 51 of the AIA, must provide also the summary of the DPIA carried out in accordance with article 35 of the GDPR, together with other information and documents[23]. The two assessments required by different regulations will therefore need to be carried out jointly in certain situations, covering different aspects and complementing each other. Not only providers of high-risk AI systems (being the subjects developing them) should have an obligation to carry out a DPIA, but also deployers, i.e. entities using the technology under their authority. Providers, while designing high-risk AI systems, must ensure that their operation is sufficiently transparent, to enable deployers to interpret the system's output and use it appropriately. These systems must also be accompanied by instructions for use, including concise, complete, correct and clear information that is relevant, accessible and comprehensible to users[24]. All these information from the providers should be used by deployers of high-risk AI systems to comply with their obligation to carry out a data protection impact assessment. Part of the information set could also be included in an extract of the DPIA carried out by the provider. It is unlikely that the provider will make available the entire assessment for these purposes, given that it may contain also confidential information or evaluations that should not become public.

23 Annex VIII of the AIA, section B (Information to be submitted by deployers of high-risk AI systems in accordance with Article 51(1b).

24 Article 13 of the AIA.

4.2. Risk management system

Pursuant to article 9 of the AIA, a risk management system shall be established, implemented, documented, and maintained in relation to high-risk AI systems and shall be understood as a continuous iterative process planned and run throughout the entire lifecycle of a high-risk AI system. The system shall comprise different steps. The first consists in the identification and analysis of the known and foreseeable risks that can occur to health, safety or fundamental rights in view of the intended purpose of the high-risk AI system. This approach is therefore similar to the one provided by article 35 of the GDPR. The main difference is that:

- according to the AIA, the situations where it is necessary to carry out the preliminary assessment are already defined (given that the high-risk AI systems are listed in the relevant annex);
- the GDPR leaves to the data controller to establish if the processing activity is likely to result in a high risk to the rights and freedoms of natural persons. The AIA, further to the initial mentioned risks analysis, requires also (always pursuant to article 9) an evaluation of other possibly arising risks based on the analysis of data gathered from a post-market monitoring system and the adoption of suitable risk management measures.

The risk management measures shall be such that risks are minimized effectively. In identifying the most appropriate risk management measures, the following shall be ensured:

1. elimination or reduction of risks as far as possible through adequate design and development of the high-risk AI system;
2. where appropriate, implementation of adequate mitigation and control measures in relation to risks that cannot be eliminated;
3. provision of adequate information as regards the risks and, where appropriate, training to deployers.

In eliminating or reducing risks related to the use of the high-risk AI system, due consideration shall be given to the technical knowledge, experience, education, training to be expected by the deployer and the presumable context in which the system is intended to be used: the overall assessment must also cover such aspects. Another requirement is that high-risk AI systems shall be tested to identify the most appropriate and targeted risk management measures. This testing shall ensure that such systems perform consistently for their intended purpose and they are in compliance with the relevant provisions set out in the AIA.

4.3. Transparency obligations as a guide to carry out a DPIA

Deployers of high-risk AI systems shall use the information provided under article 13 of the AIA to comply with their obligation to carry out a DPIA[25]. Article 13 (mirroring the relevant provision of the GDPR on privacy notices) concerns the matter of transparency and precises the concise, complete, and correct information that must be provided to users of high-risk AI systems.

From the mentioned provisions it appears, therefore, that the following elements must necessarily be considered, in addition to those provided by the GDPR, in the performance of a DPIA on high-risk AI systems (it shall be reminded that the AIA is without prejudice and complements the GDPR):

a. the identity and the contact details of the provider and, where applicable, of its authorised representative;
b. the characteristics, capabilities and limitations of performance of the high-risk AI system, including:
   - its intended purpose;
   - the level of accuracy, including its metrics, robustness and cybersecurity against which the high-risk AI system has been tested and validated and which can be expected, and any known and foreseeable circumstances that may have an impact on that expected level of accuracy, robustness and cybersecurity;
   - any known or foreseeable circumstance, related to the use of the high-risk AI system in accordance with its intended purpose or under conditions of reasonably foreseeable misuse, which may lead to risks to the health and safety or fundamental rights;
   - where applicable, the technical capabilities and characteristics of the AI system to provide information that is relevant to explain its output;
   - when appropriate, its performance regarding specific persons or groups of persons on which the system is intended to be used;
   - when appropriate, specifications for the input data, or any other relevant information in terms of the training, validation and testing data sets used, taking into account the intended purpose of the AI system;
   - where appropriate, information to enable deployers to interpret the system's output and use it appropriately;
c. the changes to the high-risk AI system and its performance which have been pre-determined by the provider at the moment of the initial conformity assessment, if any;
d. human oversight measures, including the technical measures put in place to facilitate the interpretation of the outputs of AI systems by the deployers;
e. the computational and hardware resources needed, the expected lifetime of the high-risk AI system and any necessary maintenance and care measures,

25 Article 29 of the AIA.

including their frequency, to ensure the proper functioning of that AI system, including as regards software updates;

f. where relevant, a description of the mechanisms included within the AI system that allows users to properly collect, store and interpret the logs.

4.4. Technical documentation

Article 11 of the AIA requires that, in relation to high-risk AI systems, technical documentation shall be drawn up before that system is placed on the market or put into service and then kept up-to date. This technical documentation shall be drawn up in such a way to demonstrate that the high-risk AI system complies with the requirements set out in Chapter 2 of the AIA (including also the above-mentioned provisions on DPIA and transparency obligations) and provide competent authorities with all the necessary information to assess the compliance of the AI system with those requirements.

The entire risk assessment procedure and related documentation must be formalized so that it is possible to demonstrate that AI systems (and related processing activities) are developed and used in accordance with the applicable rules.

4.5. EU declaration of conformity

For high-risk AI systems, the providers shall draw up a EU declaration of conformity that must be kept at the disposal of the national competent authorities for 10 years after the AI high-risk system has been placed on the market or put into service.

This declaration shall contain several information, also including a statement that the AI system complies with the GDPR, where the AI system involves the processing of personal data[26].

4.6. Incidents reporting

Providers of high-risk AI systems placed on the Union market shall report any serious incident to the market surveillance authorities of the Member States where that incident occurred. This notification shall be made immediately after the provider has established a causal link between the AI system and the serious incident or the reasonable likelihood of such a link, and, in any event, not later than 15 days after the provider or, where applicable, the deployer, becomes aware of the serious incident. Following the reporting of a serious incident, the provider shall, without delay, perform the necessary investigations in relation to the serious incident and the AI system concerned. This shall include a risk assessment of the incident and corrective action. The provider shall co-operate

26 Annex V of the AIA.

with the competent authorities and where relevant with the notified body concerned during the investigations.

This approach is very similar to the one introduced by the GDPR, that requires data controllers to notify a data breach to the competent supervisory authority without undue delay and, where feasible, not later than 72 hours after having become aware of it. The notification, according to article 33 of the GDPR, shall always be made, unless the personal data breach is unlikely to result in a risk to the rights and freedoms of natural persons. To evaluate this circumstance, a risk assessment must be carried out, as now required by the AIA.

5. Conclusion

While the potential benefits of Artificial Intelligence are undeniable, particularly in fields like medicine, its development and application pose significant challenges. These risks, if left unaddressed, can exacerbate existing societal issues such as discrimination and inequality. Therefore, it's crucial for developers and users of AI products and solutions to be fully aware of these challenges, as they can have profound and direct impacts on people's lives.

To mitigate these risks, proactive evaluation is essential. The AIA provides valuable guidance for organizations aiming to responsibly integrate AI into their institutional, business, and research activities. However, the GDPR remains the cornerstone of data protection within the European Union. It mandates business operators and legal tech professionals to conduct thorough risk assessments, focusing on privacy and data protection implications associated with AI systems.

Importantly, the AIA does not render existing data protection assessment methodologies obsolete. While addressing privacy concerns is crucial, a broader evaluation, encompassing other fundamental rights, is now necessary, especially when dealing with high-risk AI systems.

As we venture further into the AI-driven era, the collective challenge will be to ensure that technological progress does not come at the expense of ethical principles or societal well-being. The frameworks provided by the AIA and GDPR offer a pathway toward achieving this balance, but their effectiveness will ultimately depend on the commitment of all stakeholders to uphold and act upon these ethical imperatives. The future of AI holds the promise of enhancing human capabilities and addressing pressing global challenges, provided it is steered with a conscientious and principled approach.

Chapter VIII
Privacy in the Digital Age: a Look at the Transformation of the Concept

by Emanuele Brambilla*

INDEX: 1. The different shades of privacy. – 2. A new anthropology? – 3. Dignity as the cornerstone. – 4. Few suggestions on the link between privacy and dignity.

1. The different shades of privacy

The evolution of the concept of privacy has followed from the ongoing developments of the digital society and the Web[1]. The increasing importance of the Internet and the pervasiveness of artificial intelligence systems have created a new dimension in which privacy cannot be understood as the protection of a "material" thing. In a certain sense, in the physical world we can protect our data in an easier way, because it depends more on personality and on our desire to let people into our lives. On the contrary, the immateriality of the digital world allows our information to be present simultaneously on a certain number of devices, creating the sensation that it is more vulnerable and less controllable.

For example, when we buy something on the Internet or when we sign up on a website, we have to give away personal and sometimes even economic data, especially if we use a credit card. The personal data in discussion such as our name, address, date of birth, age, sex and more is precise enough to refer to a single individual and to nobody else. This type of information is treated as an asset, in the sense that we can exchange it for services and more, like in a sort of barter[2]. Since we cannot totally renounce to it, European and national

---

* Ph.D. student in Philosophy of law at the Department of Juridical Sciences "Cesare Beccaria" at the University of Milan, under the guidance of Professor Giovanni Ziccardi. He graduated in Philosophy (bachelor's and master's degrees, respectively in 2019 and 2022) at the Catholic University of the Sacred Heart, in Milan. From 2023 he is a Research Fellow at the Information Society Law Centre (ISLC).

1 «Two have been [...] the phenomena that have conditioned what some scholars have called "the fifth industrial revolution": the spread of the Internet and the increase in storage and processing capacity provided by computers [...] The Internet has brought about the change in a reality that is nurtured by the networked presence of human». Bonavita, S. (2017) 'Le ragioni dell'oblio', *Ciberspazio e diritto*, vol. 18(57), p. 86.

2 Bonavita writes about a double economic power of data: one is intrinsic, the other dependent to third party's interests. «Data, thanks to the processing methods now possible, has

regulations are very helpful in preventing or limiting negative consequences and in protecting our data from abuse.

The General Data Protection Regulation has these aims, since it is based both on a risk and on a right approach. If we do not interpret prevention as a total elimination of the risk (which is impossible) but as a minimization of it, we can better understand what the GDPR proposes in this regard.

The main measures that help to handle the risk are data minimization, pseudonymization and data protection by design and default. The first implies, among other things, that data should be "processed lawfully, fairly and in a transparent manner in relation to the data subject" and "collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes"[3].

Pseudonymization means "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"[4]. Since these two principles should be implemented before the beginning of the processing, they are related to privacy by design and default.

Indeed, article 25 asserts that the controller:

> […] shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriat technical and organizational measures, such as pseudonymization, which are designed to implement data-protection principles, such as data minimization" and "[…] shall implement appropriate technical and organizational measures for ensuring that, by default, only personal data which are necessary for each specific purpose of the processing are processed[5].

More specifically, privacy by design indicates that data minimization, pseudonymization, transparency and security have to be incorporated in the system's project and not after its release. On the other hand, privacy by default is centered on a balanced treatment of data, meaning "the amount of personal data collected, the extent of their processing, the period of their storage and their accessibility"[6]. The top two principles reveal the right-based approach[7], the latter two the risk

---

exponentially increased its economic value. It thus appears possible to assert an intrinsic economic value of the data as such, and an economic value relative to the information that this data carries as it becomes apparent to third parties». Ivi, p. 94.

3 GDPR, chapter II, article 5.

4 GDPR, chapter I, article 4.

5 GDPR, chapter IV, article 25.

6 Guidelines on Data Protection by Design and by Default (4/ 2019). Available at: https://www.edpb.europa.eu/sites/default/files/consultation/edpb_guidelines_201904_dataprotection_by_design_and_by_default.pdf.

7 «The right-based approach to data protection […] views data protection as a matter of individual rights. These rights are organized in two layers. The top layer includes the fundamental

approach[8]. These approaches are not separated or parallel, but synergetic to one another, as we said earlier.

On one side individual rights such as privacy are protected with data minimization and pseudonymization. On the other, the risk approach permits the flow of data with the resulting benefits, problems and services. In other words, risk does not (or should not) automatically entail that our information is threatened, but that it is safeguarded even if entered on digital devices.

It is curious that, although we have been writing about "privacy" by default and design, this precise term never appears in the GDPR[9]. If we read it carefully, article 25 speaks about "data protection", and not about privacy. This "shift" in the interpretation of the Regulation highlights the first and fundamental level of privacy, which is the only one that seems to be included in the GDPR[10]. Therefore, privacy refers to sensitive data that we have shared on the Internet and that we want to be protected from undesired intrusions or misuse.

This first level of privacy entails another one, that is to say the right to follow, govern and control our data wherever it is. Luciano Floridi legitimates and justifies this claim by saying that we should see an identity between the person and the data pertaining to her[11].

This means that the improper storage, manipulation or violation of personal information is damaging the physical being itself [12]. In other words, he is claiming that personal data should be safeguarded not to protect personal privacy,

---

rights to privacy and data protection, which are synergetic to other fundamental rights and principles: dignity, freedom of thought […] etc. The lower tier is constituted by the data protection rights granted to individuals by the GDPR, such as the power to consent and withdraw consent […], the right to information, access, erasure, and the right to object». Giovanni Sartor, The Impact of the GDPR on Artificial Intelligence. Available at: https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf (p. 66).

8 «The risk-based approach, rather than granting individual entitlements, focuses on creating a sustainable ecology of information, where harm is prevented by appropriate organizational and technological measures». Ibid.

9 The term "privacy" appears once in footnote 2 of premise 173 of the document, which refers to the Directive on Privacy and Electronic Communications.

10 This paper cannot take into consideration all the peculiarities of the four kinds of privacy (physical, mental, decisional and informational) as described in Floridi, L. (1999) 'Information Ethics: On the Philosophical Foundations of Computer Ethics', *Ethics and Information Technology*, vol. 1(1), pp. 37-56. We are interested in all the data that are present in the virtual world and we will address the topic in order to build a generally valid argument.

11 In the next chapter we will discuss how to interpret this identity. Floridi, L. (2005) 'The Ontological Interpretation of Information Privacy', *Ethics and Information Technology*, vol. 7(4), pp. 194-195.

12 Tavani underlines that Floridi should distinguish between a «[…] mere loss of privacy (in a descriptive sense) in contexts that are not normatively protected […] versus a violation of privacy (i.e., in a normative sense) in contexts or situations that have been established as "normatively private"». Tavani, H.T. (2008) 'Floridi's Ontological Theory of Informational Privacy: Some Implication and Challenges', *Ethics and Information Technology*, vol. 10(2), p. 162.

but individual's identity. This is a very strong statement because it places the digital identity on the same level as the personal identity, as we will address later. In this context, the expression "digital body" has a great relevance in terms of its informational power[13], since it refers to all our data circulating on digital devices, even those we may not be aware of. However, the right to privacy is not absolute[14], and can be violated in serious cases[15]. For example, Directive 680/2016 establishes that personal data can be gathered and processed during a criminal investigation or when the authorities suppose that a crime is about to be committed. In such situations, the Directive clearly distinguishes between the different categories of people involved (offenders, witnesses, victims, defendants), who must be informed that a processing of their data is occurring. Nevertheless, this does not imply that information can be treated indiscriminately, but that it must be stored and analyzed with fairness, transparency and for a proportionate time, based on the investigation's purposes. It is important to underline that such particular data is (or should be) managed without compromising the dignity of the subjects involved and at the same time are collected in order to protect citizens and society. A good balance between security and privacy, from one hand, cannot be reached if data is totally untouchable but, on the other hand, even if collected it should be processed bearing in mind that its misuse may affect the person to whom it refers.

2. A new anthropology?

In the previous chapter we talked about the various facets of privacy, underlining that the protection of the digital body is the most relevant of these. The aim of this chapter is to investigate the existing relationship between this body and the human person. In this regard, Floridi claims that human beings should be interpreted as "inforgs", namely as "sets of information"[16]. According to him, we live in an ecosphere that interacts with an infosphere, that is to say an "environment" made up of all the information elaborated by digital devices and ourselves. This infosphere do not simply deal with human relations, but is centered on the production, elaboration and sharing of information with artificial and human agents. Additionally, Floridi asserts that in the digital society

13 The term "digital body" appears in a lot of Rodotà's works, including: Rodotà, S. (2014) *Il mondo nella rete. Quali diritti, quali vincoli*. Laterza, Rotosud – Oricola, p. 30.

14 «The right to the protection of personal data is not an absolute right; it must be considered in relation to its function in society and be balanced against other fundamental rights, in accordance with the principle of proportionality». GDPR, premise 4.

15 We will deepen this concept in chapter three where we will discuss the differences between ontological and moral dignity.

16 He writes that we are "also" (and not only) inforgs. Floridi, L. (2020) *Il verde e il blu*, Raffaello Cortina Editore, Milano, p. 149.

personal data and individual identity are "glued together" like never before[17]. If they are so "glued" it is because we perceived a duality between ourselves and our information scattered on the Web and we wanted to correct it. This duality carries with it the unease that maybe we experienced when we entered such data on Internet pages and the urgency of their protection[18]. The claim that all our information is our "digital body" precisely represents that gluing and helps to understand the importance of data's safeguard. Nevertheless, this statement leads to various questions pertaining ontology and self's identity.

We suggest that arguing that we ontologically correspond to our data is misleading since it does capture all the depths of the human being. At the beginning of the first chapter, we said that few critical information is enough to identify a single person. However, "to identify" does not mean "to know" or "to understand" that person. It is true that the videos that we watch, the Web pages that we visit, the products that we buy contribute to define us more precisely, but even in these cases we cannot reduce the human being to that. Unfortunately, this data is useful for profiling and for statistical and economic purposes, but that does not affect the ontology. Furthermore, we cannot be sure that this kind of data really reflects the reality of the person it refers to[19]. Floridi himself seems to deny a radical ontological position preferring a relational one, although in some passages he leans toward another idea[20].

For example, in the article The Ontological Interpretation of Information Privacy he states:

> Looking at the nature of a person as being constituted by that person's information allows one to understand the right to informational privacy as a right to personal immunity from unknown, undesired or unintentional changes in one's own identity as an informational entity […] "You are your

17 «La nostra identità e i nostri dati personali non sono mai stati incollati insieme così indistinguibilmente come accade oggi, allorché si parla [GDPR, article 4] di identità personale dei soggetti interessati». Floridi, L. (2022) *Etica dell'intelligenza artificiale*, Raffaello Cortina Editore, Azzate, p. 27

18 Entering data on platforms creates a sensation similar to that which we experience when we regret telling someone a secret. This similarity concerns only the sensation that a part of us is "out there"; the feeling of regret for the revelation of the secret is caused by the sensation that a part of us is now known by somebody else, who maybe we do not totally trust.

19 An issue connected with this is the one of fake profiles. For example, on social networks we can (willingly or not, and for different purposes) invent false hobbies, opinions, personality traits and more. These characteristics should not be considered parts of our digital body since they do not truly represent us.

20 «[…] dobbiamo fare attenzione ad accettare non una posizione ontologica e assoluta, bensì una epistemologica e relazionale». Floridi, L. (2020) *Il verde e il blu*, Raffaello Cortina Editore, Milano, p. 137.

information", so anything done to your information is done to you[21].

Unlike what we have seen before in *Il verde e il blu*, Floridi in this article talks about the informational nature of the person, appearing to endorse a strict ontological view[22]. Anyway, we believe that this position does not reflect Floridi's recent thought and therefore should be softened with the following suggestions.

First of all, we should divide data in intrinsic and extrinsic to the self. On one hand, our name, surname, age, physical description and all the other personal characteristics contribute to intrinsically define our identity[23]. On the other hand, our credit card or cellphone number and our address even though have the ability to identify us, are not strictly included in our personal identity[24].

Above we called all the data that the computer "knows" our digital body. However, considered what we have just explained we can affirm that this body cannot be tautologically identified with our personal identity. If we overlay our identity and our digital body, there will always be something that exceeds[25]. The reason is the fact that the digital body has a dualistic nature. On one side is a mere sum of intrinsic and extrinsic data, on the other it contains implicit information that can be discovered thanks to methods of data combination and extraction. On the contrary, our identity is more than a mere sum of personal intrinsic aspects, since it is connected with human nature[26]. Therefore, the virtual body is the true set of information (an "inforg" to use Floridi's language) in the ontological sense[27]. Since this set includes both intrinsic and extrinsic data, it is jurisdictionally relevant and it cannot exclude elements that, if violated, could affect our personal and public life.

---

21 Floridi, L. (2005) 'The Ontological Interpretation of Information Privacy', *Ethics and Information Technology*, vol. 7(4), p. 195.

22 We can hypothesize an evolution of his thought since *The Ontological Interpretation of Information Privacy* was published in 2005 and *Il verde e il blu* in 2020 or maybe that he is talking from the point of view of a certain level of abstraction (even if he does not make it explicit and these levels must be consistent one another). See Floridi, L. (2008) 'The Method of Levels of Abstraction', *Mind and Machines*, vol. 18, pp. 303-329.

23 Of course, we get old and our appearance change during time, but our identity persists. Unfortunately, in this context we cannot deepen this argument and consequently we only concentrate on certain aspects.

24 An example is the passwords that we use. They are clearly not part of our identity, but a sort of "door" to our data. Therefore, we cannot justify their protection with the argument of the digital body since we do not protect them for themselves but for what they conceal.

25 Identity is something qualitatively, the digital body is something quantitatively and therefore a simplification of the identity.

26 The nature of the human being is something that pertains to every human, the identity unrepeatably defines a person, even if it is based on nature. However, we do not have the space to analyze both concepts, so we will only consider the second.

27 There should be a difference between the protection of a single data and the one of the digital body. The latter is the set of all our information and so it should be protected with a special attention since its infringement can cause more damage to the person this body refers to.

Let's now examine the relational interpretation of our identity as presented in *Il verde e il blu*. As we have already said, from this point of view, we can also be seen as sets of data in an informational relationship directed towards people and digital devices. Since we have argued that we do not correspond to our data, what is crucial to understand now is the relation that takes place between ourselves and digital devices, and the power of identification that data has in relation to us.

For instance, when we video-chat with someone, participate to a live stream conference or when we enter our data on websites, we virtually "travel" remaining in the same exact spot. While this may seem obvious, if we go beyond a superficial reading, we can better realize the consequences of it. Claiming that we are alone in front of our computer implies that the real proximity to other people disappears, leaving the place to a mediation produced by the computer's screen. Digital devices are always handy, but when we are using them, they "vanish" in a certain sense, since we are focused on the actions we are doing[28]. Paradoxically, our digital society while potentially providing more relational opportunities, makes us lonelier and more isolated[29]. This solitude comes from two interrelated elements: our interactions with artificial agents, which we need but that do not need us, and the mediated relationship that we have with other people, whom we need and who need us. On the other side, the relation of identification between us and our information helps us to equate the protection of privacy to the defense of the identity, without ontologically equating data to the person they refer to. If we consider personal data as traces that we leave in the virtual world and place dignity at the center of our reflections, we will be able to justify data protection in a proper way.

In this perspective, we can assume that digital devices, certain types of companies and social media with their capacities to manage and use data, implicitly reduce us to our information. This is the case when, for example, some web pages require our consent to access their content (without the possibility to decline) or when home devices use data to assist us in many ways.

What we have said so far, clarifies some of the anthropological changes brought by the increasing development of the digitalization. Among these changes, we have discussed about privacy because data nowadays has a strong influence on society, reputation, economy, personal life, and so on. Privacy certainly means data protection, but, above all, it refers to the relation between

28 «Il mondo virtuale, invece, assorbe la vista e a volte l'udito […]. Il pc […] quando lo si usa e si è collegati alla rete, scompare, così come il paio di occhiali che indossiamo smettono di esistere mentre si è concentrati sulle parole che scriviamo o sul paesaggio che vediamo». Pessina, A. (2016) *L'io insoddisfatto*, Vita e Pensiero, Milano, p. 92.

29 «Anche quando il mezzo digitale mette in contatto persone che già si conoscono […] tutto ciò avviene all'insegna della mancanza […] siamo sempre più insieme ma soli. […] l'io sembra sperimentare un potere relazionale sconfinato» but «[…] basta un blackout, una batteria scarica, la perdita di connessione per stravolgere ogni rapporto». Ivi, pp. 96-98.

human being and devices, in the sense that we should decide whether and when to exclude this dimension from our life. These challenges call us to assume a stronger responsibility and accountability since new possibilities of good and evil arise[30]. This commitment should be both intellectual and practical, individual and legislative.

We created the digital sphere and we have the power and the duty to properly govern it, avoiding undesirable consequences. As we have mentioned, a key and basic role should be assigned to dignity, on which we must built the other bricks of our reflections. Without this concept we will never be able to understand the true reason to protect data or to avoid its flow. Secondly, we should consider computers as means and not goals, and relate to digital devices in a proper way recognizing their value, but also their limits.

3. Dignity as the cornerstone

We have argued that in our digital society data protection (and even more the protection of our digital body) is one of the means to safeguard our intrinsic and extrinsic identity. Nevertheless, this statement does not automatically justify itself. Therefore, now we need to understand and explain why the real reason behind this protection human dignity is.

On this regard, the GDPR mentions the term "dignity" once in article 88, which establishes that Member States should protect employee's data with rules including "suitable and specific measures to safeguard the data subject's human dignity"[31]. The problem here is that dignity is mentioned in a very specific article and not as a basic premise. Moreover, stating that the rules that protect employee's data must embrace the protection of dignity creates ambiguities since dignity is the standard on which all the rules must be built and directed. Hence, it is crucial to reflect philosophically on dignity alone, regardless of how it is substantiated in regulations or declarations.

First of all, we should consider dignity as founded upon human nature which, simplifying, is the basic and common characteristics that all human beings share[32]. Precisely because it is based on human nature, it is wrong to interpret dignity as the so called "theory of performance" does.

This theory grants dignity to human beings only if they actively display certain rational activities and functions. In this category falls the modern and post-modern concept of dignity. The first connects dignity to self-determination

30 This responsibility derives from the fact that the virtual world «[…] possiede la capacità di diventare in sé attuale producendo comunque e sempre effetti attuali-reali sulla vita dell'io». Pessina, A., op. cit., p. 92.

31 This article is entitled "Processing in the context of employment". GDPR, article 88.

32 To deepen the knowledge of the concept of human nature we recommend: Lodovici, G.S. (2017) *La socialità del bene*, ETS, Pisa, pp. 149-160.

and liberty, the second underlines the social relevance of it and its connections with human rights[33].

The serious problem of this view is that it unacceptably denies dignity to those who do not show *in actu* rationality like children, people with certain disabilities or in a vegetative state[34]. To resolve this difficulty, we need to concentrate not on the act itself, but on potentiality, that is the ability to act or think rationally thanks to the common human nature. This connection with human nature is fundamental because it grants dignity to people with lifelong mental illnesses and refutes critics who hold that they also lack rationality in potency[35]. The true core of the argument is that the human being is a "person", namely an "individual substance of a rational nature"[36]. Specifically, every human being is a person regardless of whether he is capable or not to actively exercise rationality or relational abilities. This is something that pertains to the ontological order and so it is not affected by passivity or inaction, and it can never be lost[37]. Therefore, every person has an ontological dignity that no one can take away, since it derives from the possession of the rational human nature and not from certain capacities, laws or from the recognition of others[38].

Regarding this, Floridi argues that, after Copernico's, Darwin's, Freud's and Turing's discoveries we can no longer justify human dignity with a classic "anthropological exceptionalism"[39]. On the contrary, he interprets dignity as a common deficiency or fragility (literally, as a "work in progress"), and the human being as the exceptional "glitch" of the universe[40]. In his opinion, this is the only way to defend privacy through dignity since the other four anthropologies are inadequate. In particular, he considers the Christian philosophy "irrelevant" because it connects human dignity to God and his creation[41].

---

33 For further details see: Turco, G. (2015). ll bivio della dignità umana e la questione dei diritti. *Derecho Publico Iberoamericano*, vol. 7, pp. 20-27.

34 The ontological dignity «non dipende né dal volere in atto, né dalla consapevolezza psichica, né da alcun riflesso emotivo. Tanto meno essa risulta dallo stadio di sviluppo di organi o funzioni. In quanto tale, essa è ontologicamente originaria e costitutiva». Turco, G., op. cit., p. 38.

35 Lodovici, G.S., op. cit., p. 62.

36 Aquinas, T. *Summa Theologiae*, I, q. 29, a. 1.

37 «Sono già persone lo zigote, l'embrione, il feto, il neonato, il bambino […] il moribondo, l'anziano […] il soggetto in coma». Aramini, M. (2006). *Manuale di bioetica per tutti*, Paoline, Milano, pp. 104-105.

38 «Il soggetto umano, per se stesso, è razionale e libero, ovvero dotato di intelletto e volontà, a prescindere dal suo stadio di sviluppo, dalla funzionalità di organi e di apparati, dai progetti di vita propri o altrui. […] La dignità del soggetto non equivale alla dignità dell'atto […] la dignità ontologica dell'uomo traspare, con innegabile peculiarità, dalla concezione cristiana […] essa deriva […] dall'essere creata ad immagine di Dio». Turco, G., op. cit., p. 39.

39 He presents four exceptionalisms: the Greek and roman one, the Christian one, the one developed after the Enlightenment, the contemporary one. Among these, we are interested in the Christian one. Floridi, L. (2020) *Il verde e il blu*, Raffaello Cortina Editore, Milano, p. 151.

40 Ivi, pp. 153-154.

41 Ivi, pp. 157-158.

Differently from Floridi, we believe this philosophy is not only still possible, but also important and necessary. Moreover, even if it is true that in Catholic philosophy human dignity primarily comes from being created in God's image, Floridi's statement is incomplete. Indeed, he omits to say that being created in God's image means[42], among other things, having a rational nature and being a person[43]. In other words, dignity dwells also in human nature, which is something that it is not attributed, since it is an original possession of the human being[44]. Moreover, that does not exclude that we are fragile, since precisely on this dignity should be built the principles of common good, solidarity and subsidiarity[45]. On the other hand, ontological dignity is inextricably correlated with moral dignity, which depends on individual moral merit. The ethical dignity has some similarities with the theory of performance since it is not possessed by everyone, and it is connected with actions and their evaluation[46]. In other words, thanks to our free will we can impress goodness or wickedness to our actions, consequently creating a virtuous reinforcement to our morality or the loss of moral dignity[47]. On the contrary, ontological dignity can never be lost since it pertains to the ontology of the human being, and not to the expression of certain acts or to the legislation. Thanks to this argument, we can even justify the provisions of Directive 680/2016 that legitimate data's storing during criminal investigations. However, these investigations have to be conducted without contradicting or despising the ontological dignity of the people in question.

4. Few suggestions on the link between privacy and dignity

The point of connection between dignity and privacy in our digital environment is difficult to fully define. This is due to the fact that dignity can

42 Thomas Aquinas writes: «[…] per imaginem significatur intellectuale et arbitrio et per se potestativum». Rationality, free will and power on our actions are considered in potency, in the sense that they are possessed by the person even if she is not able to manifest them. Thomas Aquinas, Summa Theologiae, I-II, prologus.

43 To consult the definition of the term "person" see: Thomas Aquinas, Summa Theologiae, I, q. 29, a. 1-3.

44 Russo, F. (2021) *Antropologia filosofica. Persona, libertà, relazionalità*, EDUSC, Roma, pp. 18-19.a

45 Lodovici, G.S., op. cit., p. 65.

46 The point of connection between ontological dignity and the theory of performance is that since the first is possessed by everyone it must be recognized by all human being and by the law. The fact that this dignity is not recognized does not change its consistency and reality. «[…] il riconoscimento [of dignity] è dovuto, è un dovere morale di giustizia […]. Il pieno adempimento di questo dovere di giustizia richiede che esso si manifesti anche nelle forme del diritto e della politica». Viola, F. (2017) 'Dignità umana, diritti e legge naturale', *Prudentia Iuris*, vol. 83, p. 39.

47 Aristotele, *Etica Nicomachea*, II, 5, 1103a 14 – 1103b 7 e II, 5, 1105b 19 – 1106a 13.

theoretically underpin different and conflicting rights at the same time, making it complicated to reflect without falling into contradictions. Hence, we will now propose some suggestions, without claiming to have been exhaustive.

First of all, founding privacy on dignity emphasizes that individuals have the right to maintain control over their personal information. This ensures self's autonomy, allowing people to decide what data they want to share and with whom. It also means that they should not be obliged to consent to the processing of their information if they do not want to be excluded from the access to a particular service[48].

This is a fundamental aspect of privacy, which requires people to be informed about how their data will be used and where it will be stored. This principle respects dignity by acknowledging people's right to take informed decisions about their personal information. However, that does not imply that one should be completely free to treat his personal data indiscriminately, since there are ways in which people use their data that are contrary their intrinsic dignity.

One way to avoid this is to educate Internet users on how to protect and use their information, by raising awareness of the risks on the Web: providing them with a clear ethical foundation could help to prevent them from the impairment of their dignity.

A second possible link of the two concepts, is that dignity helps privacy to protect citizens from surveillance and monitoring, that is to say, from being treated as mere "objects" of observation[49]. In other words, dignity is crucial for security and safety, offering protection against numerous threats such as identity's theft or storage, illegal profiling, harassment or cyberbullying, and creating a safe space for personal and social interactions.

A third suggestion, connected to the second, is that dignity helps to mitigate the amount of personal data that can be used to unfairly target and exclude individuals. This allows to defend users from reputational damages[50], that can unfairly affect their personal and public lives.

In conclusion, philosophical and moral guidelines, normative frameworks, and appropriate technological measures can all play a role in safeguarding human dignity in the digital age, but only if we start from a concept of dignity that respects the uniqueness of the human being.

48 GDPR, chapter II, article 7.

49 Anyway, this statement does not exclude what we have mentioned about Directive 680/2016. To further explore the concept of the so called "society of surveillance" read: Barberis, M. (2017) *Non c'è sicurezza senza libertà*. Il Mulino, Bologna, pp. 88-111.

50 See Bonavita, S., Poli, V. (2017) 'La tutela civilistica della reputazione online', *Ciberspazio e diritto*, vol. 18(58), pp. 307-318.

Chapter IX
The Regulation of Data Brokers in Europe: How to Address an International Data Governance Challenge

by Isabela Maria Rosal*

INDEX: 1. Introduction. – 2. The market of data brokers. – 3. Possible good (regulatory) news: the Data Governance Act. – 4. Beyond the DGA: more regulatory limits for data brokers. – 5. Conclusions.

1. Introduction

Control over your data is protected as a fundamental right since exploring this information can lead to severe consequences such as manipulation, anti-competitive practices, misuse of one's identity, and invasion of privacy, beyond various others. Nonetheless, there are critical economic effects from processing personal data, including innovation, more competition in a concentrated market, personalized products, and more efficiency, among other positive consequences. Therefore, it is imperative to find a regulatory balance to protect individuals and collective negative effects that may emerge from data use, while also allowing economic development, forming a positive data ecosystem. For this, regulators and law actors must consider the different risks related to data processing activities.

One important agent that still acts without further consideration of the effects on data subjects is the actor commonly known as the "data broker". Data brokers are a part of the data ecosystem, which includes all data, all transactions, and the spaces involving personal data processing activities. The complexity of data ecosystems and the protection of personal data are not incompatible with the existence of agents handling personal information. However, the data processing activities should always observe the rules that guarantee data and privacy protection, what is only possible if the data subjects understand the other agents involved in this ecosystem[1]. Brokers add to the complexity of

* Legal doctoral researcher at imec – KU Leuven – CiTiP, holding a Master's law degree from the University of Brasília. Her research focuses on data protection, technology regulation, cybersecurity, privacy and competition.
ORCID: https://orcid.org/0000-0003-1604-7105. E-mail: isabelamrosal@gmail.com.

this system by working in regulatory gaps, leading to financial transactions of sets of personal data to various economic agents globally, without effectively informing the data subjects. Therefore, these actors follow a business model that generates more insecurity in data subjects, who have their vulnerabilities explored and lose control over who has access to their personal data.

This work tackles the problem regarding data brokers and their transnational existence. Although they work in a market that sees limitations regarding the fundamental right to data protection and privacy, these agents can find loopholes in the regulatory frameworks and then continue working and profiting from personal data exploration in different markets and countries. This led to public agencies and private actors having access to several types of data – what can include sensitive data. Regulating these agents is even more difficult considering the transnational approach of the companies, building a scenario that makes it seem almost impossible to move forward without international cooperation and intervention. Additionally, the sources of information used by brokers are broad and not well mapped. Several times, the data broker does not even know how the data aggregators collect information, bringing the possibility that they got data from infinite sources. From this scenario, the paper aims to understand how the Data Governance Act (DGA) brings light to the discussion since it is a step forward on regulating the so-called "data sharing services", including data intermediary services[2].

As a European Regulation, the DGA may serve as an example of how to tackle this global challenge. The norms establish further rules regarding data management, including detailed transparency obligations. However, it is still unclear if data brokers will fall under the scope of the DGA. Thus, this study intends to highlight how classifying these agents as data intermediation services may be a solution for limiting the power of brokers.

The research also analyses which topics need further addressing, especially third-country data transfers and higher level of protection to special categories of data. The conclusions will also nod to the debate about inferred data and anonymisation, as one must consider that data brokers work with different data sets that initially represent only non-personal data but the aggregation of information might lead to the identification and profiling of subjects. For these goals, the paper is divided into three topics. The first consists of an initial definition of the market of data brokers, with criticisms regarding their approach to exploring personal data. The second topic presents a brief defence of how the DGA may be used as a regulatory framework to bring limits and enforcement to data brokers. The last chapter consists of a summary of various topics that

1 World Economic Forum. (2022). Advancing Digital Agency: The Power of Data Intermediaries.

2 See https://www.euractiv.com/section/digital/news/new-eu-data-brokers-wont-have-to-be-european-commission-says/.

must be object of different regulations to limit the power of brokers and guaranteeing better protection of personal data.

2. The market of data brokers

Data brokers are commonly known as "companies that collect consumers' personal information and resell or share that information with others"[3]. However, no formal and comprehensive definition describes these agents well: some consider that the exploitation of data for monetization provided by other companies is enough for a company to be classified as a data broker while others understand that, for this definition, the company cannot be the own source of data – what would exclude some organizations such as Facebook from the concept, since it collects data mainly from its own products[4]. Anyhow, it is widely known that their activities start with the massive collection of data – personal or not –, followed by the processing of the data, which leads to the monetization of the information by brokers since they sell this knowledge of individuals to different companies, or even public bodies, constructing digital profiles of individuals with direct effects, such as the credit scores used by determined someone's access to financial loans. Sources of data vary from different regions and agents involved. To collect data, brokers may buy data sets, scrape public records, data made publicly available by data subjects, or other accessible sources of information. For this, they might use third-party tracking technologies. However, contrary to self-determination, these agents work in secrecy. So, data subjects do not know how their information is processed, having to deal with the consequences of a digital profile that they have no control over[5]. Effects are even more concerning when one considers that after collecting information from various sources, the brokers cross the data with other information previously collected by these actors. This process leads to new inferred data (the aggregation effect), leaving the subjects lacking control of their information. Beyond this, the inferred data can be inaccurate or outdated, or even real vulnerabilities that should be private[6]. Additionally, data brokers are not consumer-facing companies, which brings obstacles to applying transparency rules and building trust in these agents that still work in the shadows of

3 Federal Trade Commission (2014). Data Brokers: A Call for Transparency and Accountability.

4 Reviglio, U. (2022) 'The untamed and discreet role of data brokers in surveillance capitalism: a transnational and interdisciplinary overview', *Internet Policy Review: Journal on internet regulation,* vol.11(3).

5 Rodotà explains that personal data processing in the digital world becomes a form of a digital manifestation of an individual. The digital profile created is one of the main reasons one should have the guarantee of control over the use of their data. (Rodotà, 2008)

6 Mishra, S. (2021) 'The dark industry of data brokers: need for regulation?', *International Journal of Law and Information Technology*, vol. 29, pp. 395-410. Available at: https://doi.org/10.1093/ijlit/eaab012.

regulatory gaps. Also, the complexity of the data ecosystem makes it difficult to guarantee that data subjects receive accurate information. For example, after cross-references, it becomes more difficult to trace the exact sources of data.

Another obstacle is the protection of the brokers activities by companies' secrecy, especially the processes of creating inferred data[7].

Another aspect that allows the continuing work without clear boundaries of data brokers is the fact that their activities do bring advantages to the consumers and to the economy. Considering that data has economic implications, data brokers do allow that different providers have access to information that may be economically essential for the development of their activities. Nonetheless, it must be considered that most activities of brokers are non-compliant to specific data protection law, which allows big data brokers to exploit consumers that are in a asymmetric relationship with these agents[8] (Nie & Han, 2018).

But these aspects cannot be a complete stop to the regulation process of the activities developed by the brokers. Even though there are cases involving oversight (and subsequently sanctioning) of brokers' activities, these agents are very well established in the market, even without observing data protection rules (e.g., General Data Protection Regulation – GDPR). So, the existing regulatory framework, including the GDPR, does not effectively regulate the data ecosystem. While data protection law empowers individuals, it also brings several burdens for individuals to follow their data in a very complex system of data processing. Thus, limiting data brokers' actions relies on new regulatory initiatives that consider the limits of individuals in parallel with the existing data economy. Following the recent developments in Europe, the new Data Governance Act may address the mentioned challenges.

3. Possible good (regulatory) news: the Data Governance Act

Consumers' trust in companies is essential for the data economy. By the assumption that organizations will process personal data lawfully and proportionally, subjects are more engaged and allow broader efficient processing of their data[9]. Consumer-company trust is imperative for the economic development of the European Single Market. For this scenario, transparency and accountability practices must be part of an organization's daily activities without

7 Crain, M. (2019) 'The limits of transparency: Data brokers and commodification', *New media & society,* vol. 20(I), pp. 99-104.

8 Nie, Y., & Han, X. (2018) 'Research on consumers' protection in advantageous operation of big data brokers', *Cluster Computing.*

9 Jai, T.-M., King, N. J. (2016) 'Privacy versus reward: Do loyalty programs increase consumers' willingness to share personal information with third-party advertisers and data brokers?', *Journal of Retailing and Consumer Services*, pp. 296-303.

opaque exploration of vulnerabilities. As mentioned, the complexity of the data ecosystem in a digital economy calls for the action of new players. For that, the policy-makers must understand how the system runs and try to find possible and more concrete answers on how to protect the data subjects' rights. As pointed out above, data protection law was not enough to avoid obscure actions from data brokers, which calls for new policy actions.

In this scenario, the new European proposals part of the European Strategy for data occupy a relevant space[10], especially the Data Governance Act (DGA)[11], which sets different obligations to agents that work on the data ecosystem. The new regulation applies to all "services which aim to establish commercial relationships for the purpose of data sharing"[12]. Recital 3 of the DGA acknowledges the need for improvement of the conditions for data sharing in the internal market to guarantee the existence of data flows, but with a human-centric approach. As the DGA foresees the possibility of data intermediation, which is defined as "a service which aims to establish commercial relationships for the purposes of data sharing between an undetermined number of data subjects and data holders on the one hand and data users on the other", it is still essential to understand if this would apply to all data brokers into a level/definition of data intermediaries, or if they would be considered as a different agent in the ecosystem. This new definition of data intermediaries would allow data subjects to delegate the control of their data to an organization that would identify the different agents responsible for the processing of data, while guaranteeing the respect to the subject's wishes[13]. This brings a lot of responsibility to the data intermediaries, since they must understand the whole ecosystem in which they are using the data, so they can rely work as an advocate to the data subject and the organizations interested in the information. So, it is already clear that the new provisions move the regulatory scheme from general rules about transparency to a system more action-focused, bringing more trust from the subjects to the data systems[14]. From this, the most human-centric interpretation would be that the main services offered by data brokers should follow the rules of

10 European Commission. (2020). A European Strategy for Data, Brussels.

11 Full text of Regulation (EU) 2022/868 is available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022R0868&from=EN.

12 Recital 28 of the Regulation (EU) 2022/868 also states that the Regulation only applies to "directly concern the provision of data intermediation services", what may allow brokers to still find regulatory gaps to continue offering their services in a non-compliant manner, the broad understanding of intermediation services seems to cover the main activities developed by brokers.

13 World Economic Forum. (2022). Advancing Digital Agency: The Power of Data Intermediaries.

14 Recital 5 of the DGA stablishes that "Action at Union level is necessary to increase trust in data sharing by establishing appropriate mechanisms for control by data subjects and data holders over data that relates to them, and in order to address other barriers to a well-functioning and competitive data-driven economy".

the Regulation. The construction of new regulatory provisions should bring possibilities that would allow the real and effective control of the data by the owners, even if indirectly. From that, the relationship that now is disbalanced can find a way to mitigate the asymmetry between data processors/controllers and individuals.

It is important to notice that the DGA already excludes from its scope[15]:

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

So, it is possible that data brokers' companies try to mitigate the need to comply with the Regulation, which may lead to the need of new norms that would clarify what activities follow under the scope of the DGA. The European Data Innovation Board[16] has an essential role to guarantee that the data brokers are acting in a way to supply the benefits of their activities while still complying to the individuals' rights and can be the player that also confirms that data brokers should follow the DGA as intermediation services.

Even though in literature the idea of data brokers being considered as intermediation services is broadly accepted, it is important to notice that the Policy Report "Mapping the landscape of data intermediaries", published by the Joint Research Centre, the European Commission's science and knowledge service, adopts a very cautious and restrict interpretation of the intermediation services foreseen in the DGA. By explaining that the model of intermediaries presented in the regulation may not be compatible with profit models for private companies and that brokers will not be under the scope of the DGA "in case their goal is not to establish commercial relationship between data subjects/holders and data users, or whether they aggregate, enrich or transform data with aim of adding value and licencing its use to data users"[17], the study distances the

15 Article 11 of Regulation (EU) 2022/868.

16 Article 30 of Regulation (EU) 2022/868 establishes the tasks of this new European body.

17 European Commission, JCR Science for Policy Report (2023). Mapping the landscape of data intermediaries: Emerging models for more inclusive data governance. Luxembourg:

rules established by the Act instead of imposing them into brokers and better regulating this market. Although the points made by the Report are important, in accordance with the Regulation[18], and compatible to the economical possibilities behind the models[19], it is crucial that policymakers adopt more clear guidelines on the rules on the regulation of brokers. In case this is not done, the agents will continue to fall outside of regulatory frameworks and will maintain their business model that is incompatible with the fundamental rights and data protection, not notifying themselves as data sharing services.

4. Beyond the DGA: more regulatory limits for data brokers

Even if one considers that the DGA regulates the data brokers, there a still some topics that deserve more normative attention to guarantee that these agents comply with the ideas behind the fundamental right to data protection and privacy. In this study, beyond the proposition of having the DGA as a regulatory protection that establishes limits to the activities of brokers, agenda points for the better regulation of these agents are presented below, with focus on the different types of data handled by the data brokers. These points are not exhaustive and should be read as a starting point of the continuous work on the data protection field. Anonymised data does not fall under the scope of relevant regulations such as the GDPR[20]. Therefore, one of the strategies of data exploration applied by data brokers is collecting data that initially represents anonymous data. However, it is possible to infer personal information from apparent non-personal data or even data of other persons (via aggregation, enriching or transforming data). Inferring personal data by anonymised datasets

Publications Office of the European Union.

18 Recital 28 of the DGA establishes: "This would exclude services that obtain data from data holders and aggregate, enrich or transform the data for the purpose of adding substantial value to it and license the use of the resulting data to data users, without establishing a commercial relationship between data holders and data users. This would also exclude services that are exclusively used by one data holder in order to enable the use of the data held by that data holder, or that are used by multiple legal persons in a closed group, including supplier or customer relationships or collaborations established by contract, in particular those that have as a main objective to ensure the functionalities of objects and devices connected to the Internet of Things".

19 On this, see CiTiP.

20 Recital 26 of the GDPR clearly states that "The principles of data protection should therefore not apply to anonymous information, namely information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the data subject is no longer identifiable. This Regulation does not therefore concern the processing of such anonymous information, including for statistical or research purposes".

is possible since data is contextual[21]. This system brings some difficulties in guaranteeing that anonymised data should not have the same level of protection as personal data and implies challenges in guarantying the anonymization of information. Anonymous data outside of GDPR's scope leaves an interesting regulatory gap for data brokers: they can work with anonymised datasets that can still reveal personal information for the companies' owners of other datasets. This possibility means that a company that owns an established dataset with information from various subjects may acquire an anonymous dataset and have more inferences on the identified individuals. Therefore, regulations must consider the full impacts of data – personal or non-personal – exploration by a contextual approach.

In this context, the proposal of a "right to reasonable inference"[22] is not enough, since it does no bring direct obligations to the brokers who work with anonymised data. This idea establishes that every person is entitled to have some level of accuracy in the aggregated, enriched or transformed data. Nonetheless, it seems more reasonable and in accordance with the EU data strategy and case-law that as long as the inference may be connected to a specific person, this should be considered as personal data. However, this has also not been enough to avoid abuses from data brokers in the last years.

The DGA already brings relevant provisions on this topic, as the idea that even non-personal data should be protected from unlawful or unauthorised access[23]. Another category is brought by the new regulation, the so called extremely sensitive non-personal public data. However, these provisions are more focused on data held by public bodies, which brings some innovation for the idea of conceptual data but it is still limited. So, for a proper regulation of brokers, better instructions regarding handling of anonymised and aggregated data should be put into place. Considering that vast amounts of information are processed to guarantee the monetization of data brokers, it must be noted that the processed datasets may count directly with special categories of data. In other cases, aggregation of different datasets may lead to sensitive inferences. The discussion, however, should be further investigated, considering more than just the evaluation of which data is part of the brokers' activities, but also the buyers and contexts in which the data is applied. Considering that the inferences are sold to different agents, information sold by data brokers are acquired by agents in education industry, healthcare or insurance providers (Open Society Foundations, 2016). So, the level of protection that brokers must comply with

21 As explained by the World Economic Forum, the fact that data is contextual "means that non-personal-data may become personal in nature depending on the context, for example, if combined with other datasets". (World Economic Forum, 2022).

22 Wachter, S., Mittelstadt, B. (2019) 'A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI', *Columbia Business Law Review*, vol. 2.

23 Recital 15 of Regulation (EU) 2022/868.

should also consider the secondary uses of the data. Thus, it is crucial that norms are established to define specific limitations on possible transactions handling data, enumerating unacceptable exploitation of data.

Finally, even though it is a topic of big interest and regulatory development, data brokers still act beyond borders. The market explored by data brokers allow that third-countries with a less strict regulatory framework receive personal data from regions with an elevated level of protection. It is essential that the regulation of data brokers is enforced globally, working into more homogenous ideas and concepts of data and agents that are part of the data ecosystem[24].

5. Conclusions

Data brokers have established economically as intermediaries of the data ecosystem. However, these agents have found possibilities of continuing with their activities regardless of novelty regulations establishing limits and principles for data (personal and non-personal) protection. This short study defends the idea that the DGA should be used as a regulatory source to impose rules safeguarding individuals for the data brokers. Even though this Act does not bring rules for all the issues and risks related to the data ecosystem, having the data brokers falling under the definition of "data intermediation services" may bring more security for individuals, with possibility of development of new business models in the data society.

Additionally, the paper also enumerates some of the topics that must be better addressed by policy makers in the evaluation of data brokers, especially: use (and re-use) of anonymised data; inferred data, including special categories of personal data; and third-country transfers. Currently, the complexity of the data system and data exploration is more evident. Thus, regulations can set more burdens to the economic agents rather than to the individuals, what is seen in the European Data Strategy. So, it is crucial that interpretations of the new rules follow this rationale.

Following the understanding that data protection is contextual, and information about one group of people may reveal details about a different number of individuals, the DGA should be seen as an initial effort of regulating data brokers, with the need of more homogenous and global initiatives to guarantee an effective level of protection of data holders.

24 Reviglio, U. (2022) 'The untamed and discreet role of data brokers in surveillance capitalism: a transnational and interdisciplinary overview', *Internet Policy Review: Journal on internet regulation,* vol. 11(3).

Chapter X
NFT: Privacy and Author Protection

by Marco Alagna*

INDEX: 1. Introduction. – 2. NFT and copyright. – 3. NFT and privacy. – 4. Conclusion and final remarks.

1. Introduction

The concept of NFT originally comes from a token standard of Ethereum, aiming to distinguish each token with distinguishable signs[1]. This type of token can be bound with virtual/digital properties as its unique identification. NFT is a unique token that cannot be exchanged like-for-like, making it suitable for identifying something or someone in a unique way. Once an NFT has been generated in the blockchain, it is very difficult to block its diffusion, even legitimately, especially if the identity of its owner and/or creator cannot be traced for the privacy reasons that characterize the underlying technology itself. But, how to block this diffusion?

1. Burn: "Burning" an NFT is similar to effectively destroying it. While technically NFT always remain on the blockchain, you can remove one from circulation by sending it to an inaccessible wallet address (not always with fees)[2];

---

* Marco Alagna earned his Master's Degree in Law with a thesis on Legal Informatics and Private International Law, focusing on Smart Contracts and Fintech, from Alma Mater Studiorum - University of Bologna in May 2021. His main interests focus on the intersection between technology and law. He currently practices law at the boutique law firm Perani Pozzi Associati in Milan, specializing in privacy, data protection, and civil law, with a particular emphasis on new technology law. He has authored several contributions on topics such as the right to be forgotten, sections of manuals and papers on privacy, as well as works on blockchain, NFTs, the metaverse, and new technology.

1 Wang, Q., Li, R., Chen, S. (2021) 'Non-Fungible Token (NFT): Overview, Evaluation, Opportunities and Challenges', *Axirv – Cornell University*. Available at: https://arxiv.org/abs/2105.07447.

2 Ghelani, D. (2022) 'What is non-fungible token (NFT)? A short discussion about NFT Terms used in NFT', *Authorea*. Available at: https://www.authorea.com/users/511573/articles/588778-what-is-non-fungible-token-nft-a-short-discussion-about-nft-terms-used-in-nft.

2. Delist: when removing an NFT from an exchange, but involves fees (e.g., gas fees). NFT communities often encourage holders to delist NFTs ahead of major announcements[3].

In this abstract, we will attempt to briefly explore the issues between the use of NFTs and blockchain technology and the prerogatives under the GDPR for the processing of personal data, as well as how these technologies can lead to copyright and intellectual property infringement, negating the need to protect the exclusive rights holder.

2. NFT and copyright

Regarding copyright, the biggest issues may concern digital content, which is the subject of NFTs. For example, creating an NFT representing a copyrighted work poses issues in relation to the information stored on the blockchain. There is also an issue related to the dissemination of such content. The amount of information contained in the blockchain is stored not in one device, but in a number of different devices, all connected to each other, according to the technological features of the blockchain itself. NFTs, in addition to having generated a very strong interest in the crypto market in general, have been a real revolution in the field of copyright, in particular in the field of "crypto art". Currently, according to the legislation in force in Italy[4], the author of an intellectual work protected by copyright enjoys two different types of rights: moral rights and patrimonial rights.

Moral rights give to the right holder a special form of protection, which consists of deciding what to do with his work, claiming authorship, and opposing any alteration or modification of the work itself. The patrimonial rights are the so-called "economic exploitation rights" of the work, which can be translated, indicatively, into the right to: publish, reproduce, communicate and/or distribute to the public, as well as to grant the use or in toto to third parties. It is important to note that in the context of NFTs, the person who purchases an NFT token does not acquire the copyright to the content that is the subject of the token through its use, but rather acquires a "simple certificate" that allows him or her to track and prove ownership of the digital copy purchased. The recurring question when discussing art and NFTs, or for that matter the relationship between the "rights" of the token and the underlying work, is: where are these works of art, essentially? The answer is a bit more complex than one might think and depends on the type of token being created. In most cases, if the object of the token is, for example, a media file, it can be found in a special file manager (or, in technical language, a peer-to-peer protocol) that

3 Ibid.

4 Italian Law n. 633/1941.

operates on the blockchain and makes the content accessible[5]. Of course, what makes the NFT unique is the "fingerprint" of the file itself, i.e. its *hash*, which precisely identifies it. For example, if an unlimited number of NFTs of the same item are sold to different parties, each of them will own a single copy, but not the underlying original. The latter will remain the exclusive property of the author, precisely by virtue of copyright, with all that this implies, who will then be able to economically exploit his work an unlimited number of times, being remunerated for the purchase of a token linked to it. While NFTs certainly represent a new form of protection, they raise a number of concerns regarding aspects of the uniqueness of a work of any kind. It is now common knowledge that any multimedia content can be the subject of a NFT, be it audio, video or image. There is also no doubt that NFT (blockchain) technology can facilitate the management of the "piracy" phenomenon, given the extreme ease with which digital content can be duplicated by anyone who comes into possession of it. In fact, from this perspective, blockchain and the use of NFTs certainly represent an effective tool for companies to protect their brands.

3. NFT and privacy

With regard to possible issues between blockchain (NFT) uses and privacy, the main problem is in the idea that the "blocks" placed on the blockchain are themselves protected because they are encrypted, or that the contents of these should not be considered personal data. For this reason, one could conclude that what goes into a blockchain does not fall within the scope of General Data Protection Regulation (GDPR) and other privacy regulations. Without expanding and going into detail, we define "personal data" as any information that directly or indirectly identifies a specific individual. Therefore, it would be necessary to investigate whether the hash could also identify the "owner" or "creator" of the token, in the same way as has already been analyzed for IP addresses[6]. The GDPR aims to protect fundamental privacy rights. It strives to achieve this goal by giving individuals more rights and more control over their personal data. Additionally, it puts more obligations on the data controllers' shoulders and demands that controllers can always demonstrate compliance (or accountability). The GDPR is technologically neutral, which means GDPR compliance must be ensured whenever personal data of natural persons (the data subject) are processed in a structured manner. Consequently, the material scope of the GDPR is also applicable to the blockchain whenever personal

5 See, for example, IPFS available at: https://ipfs.tech.

6 See what is reported by the ICO at: https://www.authorea.com/users/511573/articles/588778-what-is-non-fungible-token-nft-a-short-discussion-about-nft-terms-used-in-nft.

data of a data subject are processed. But what might be the "friction points" between what the GDPR requires and what applies to blockchain?

First of all, the GDPR "assumes" that there is at least one data controller which can be addressed by a data subject (an individual). In contrast to this "one-on-one" fiction of the GDPR, the blockchain works with multiple players and decentralization. This makes the allocation of responsibilities on the blockchain under the GDPR more onerous. For the second, the GDPR grants certain rights (e.g. the right to data rectification, the right to data erasure etc.) to the data subjects, which again contradict the "blockchain values", since the blockchain is a shared, immutable ledger for recording transactions, tracking assets and building trust. For the third, another obligation of a data controller is to ensure "privacy by design": simply put, this obligation demands that the GDPR principles are taken into consideration in the developing phase of a product rather than later on. In this way, article 25.1 of GDPR, requires controllers to implement appropriate safeguards: "both at the time of the determination of the means for processing and at the time of the processing itself". This means that the blockchain itself should respect GDPR principles, such as data minimization and purpose limitation, but also storage limitations[7]. Lastly, the aforementioned issue of whether it is possible to consider information on the blockchain as personal data, which, depending on the case of DLT[8], may also be public and visible to all users on the chain.

4. Conclusion and final remarks

In conclusion, given the aforementioned reasoning, more regulation of new blockchain technology-based tokens and their "compatibility" with regulatory frameworks is encouraged. Specifically, such tokens should be addressed and made compliant with the prerogatives of the GDPR, first of all, privacy by design, and national intellectual property and privacy regulations. In further, it would certainly be necessary to identify, preferably at the regulatory level, what technical and organizational security measures and contractual measures are needed to guarantee the rights of users (for privacy) of creators (for copyright).

7 See https://www.lexology.com/library/detail.aspx?g=61e0954b-c262-4d0f-b33a-2bfc4f-5cb7c7.

8 "DLT" means the Distributed Ledger Technology.

Chapter XI
Digital Inheritance in Accordance with the Right to Data Protection in the Brazilian Legal System

by Guilherme Vargas Puchta*and Zilda Mara Consalter**

INDEX: 1. Introduction. – 2. Digital assets: some legal aspects. – 3. Personal data: constitutional right and person's right. – 4. Responsibility for data protection. – 5. Current status of digital inheritance: bills, court decisions and data policy. – 6. Conclusion.

1. Introduction

In the context of the Digital Revolution, new products and platforms emerged, being capable to compose a person's property – for example, songs, movies, social profiles, e-books, domain names, e-mails, tweets, passwords, data in virtual games, cryptocurrencies, non-fungible tokens (NFTs), digital signature, and many others – compounding the digital assets collection. Therefore, in the digital era, it is fundamental to establish how the transmission of this group of assets should occur after death, which corresponds to the institution of digital inheritance. However, when discussing the digital assets inclusion in inheritance, there is a collision between two constitutional rights: the right of inheritance and the right of data protection, which is provided, respectively, in items XXX and LXXIX of the article 5 of the Brazilian Constitution. To mitigate this tension, it is important to start this essay by understanding the differences between digital assets in two categories: patrimonial and existential (there are examples in the State Courts' decisions recognizing this distinction and some that ignore it). Also, some Law Draftings that completely leave out the existence of these two groups and even legitimize the free modification of existential digital assets of a deceased person by his heirs, which evidently violates the constitutional right to data protection are pending in the Brazilian Congress. Thus, considering this gap in knowledge and in the Brazilian legal system, this investigation – which is part of the Voluntary Program for Junior Scientists (PROVIC/PROPESP) of Ponta Grossa State University – has the main objective of analyzing the establishment of digital inheritance in the Brazilian legal system. As specific objectives, this research aims to analyze the insertion of digital inheritance in the Brazilian legal system; to understand the impacts

of inserting the constitutional right to data protection in the Constitution; and to contribute to the enactment of laws to defend rights in the digital universe.

To achieve the objectives above, the approach of this research applied the deductive method, starting from general concepts involving digital assets to arrive at the specific analysis of digital inheritance and its consequences[1]. In addition, it is important to highlight that this is one theoretical research supported by indirect documentation, notably doctrinal, as well as legislation and Courts decisions. Finally, it is necessary to inform that there is no field research yet, and the results obtained until now will be discussed in the sections below.

2. Digital assets: some legal aspects

Initially, it is important to inform that according to Bevilaqua[2] an asset is something that has a utility, not necessarily economic. It is possible to define digital assets as something that can be owned in the digital universe[3], has an intangible and personal feature, brings some utility to the person with economic expression or not[4], and is stored on electronic devices or another means accessed by a contract.

Under this bias, it is possible to categorize the digital assets into two groups: those with economic value (patrimonial assets) and those without it (also called affective or existential assets). This distinction asserts that only digital patrimonial assets constitute digital inheritance and are, therefore, a reason for legal

* Undergraduate student at the State University of Ponta Grossa (UEPG). Researcher at the Lawgorithm association focused on artificial intelligence and law. Volunteer Program participant of Scientific Initiation (PROVIC/UEPG). Member of the research groups: "Theory and practice of contemporary private law" (dgp.cnpq.br/dgp/espelhogrupo/0203115420872092) and "MindTheGap" (dgp.cnpq.br/dgp/espelhogrupo/2363453077537632). Author of different publications. Orcid: https://orcid.org/0009-0000-1694-2826. E-mail address: guilhermepuchta09@gmail.com.

** PhD in Civil Law (Faculty of Law Largo de São Francisco – University of São Paulo (USP) (2016)). Master Science in Business Law (State University of Londrina (2004)). Graduated in Law (State University of Maringá (1995)). Leader of the CNPq Research Group named "Theory and practice of contemporary private law" (dgp.cnpq.br/dgp/espelhogrupo/0203115420872092), since 2007. Professor of Civil Law at State University of Ponta Grossa since 2008. Author of several books and articles. Legal reviewer. Orcid: http://orcid.org/0000-0002-4257-0939. email address: zilda@uepg.br.

1 Lakatos, E. M.; Marconi, M.A. (2011) 'Fundamentos de Metodologia Científica', *São Paulo: Atlas*, ed. 6, p. 256.

2 Bevilaqua, C. (2003) 'Teoria Geral do Direito Civil', *Campinas. SP: RED Editora*, p. 155.

3 Sherry, K. (2012) 'What Happens to Our Facebook Accounts When We Die? Probate Versus Policy and the Fate of Social-Media Assets Postmortem', *Pepperdine Law Review*, vol. 40(1), pp. 185-250. Available at: https://digitalcommons.pepperdine.edu/plr/vol40/iss1/5.

4 Zampier, B. (2021) 'Bens digitais: cibercultura, redes sociais, e-mails, músicas, livros, milhas aéreas, moedas virtuais', *São Paulo: Editora Foca*, ed. 2, pp. 63-64.

interest related to their transmission to heirs, considering the continuity of its economic activity. It should be highlighted that this differentiation of digital assets is essential to avoid violating the constitutional right of protection of personal data and person's rights.

In this perspective, for an existential digital asset to be transmitted to an inheritor, it is necessary that the deceased leaves a manifestation of will, either by testament or through a choice made in the platform where the data is located.

3. Personal data: constitutional right and person's right

Considering that the focus of research is digital inheritance in the face of data protection, it is essential to present the definition of data and the duality of right to data protection, which can be considered a civil right and was recently inserted in the Constitution as a human right (public). Initially, data can be defined as a documented observation or measurement result[5] or even as a representation of facts, concepts, or instructions in a legal way that adapts to communication, interpretation, and processing by the human being or through automatic machines[6]. In addition, based on the Brazilian Law n. 13.709 (2018), named General Data Protection Law (GDPL), personal data can be defined as information related to an identified or identifiable natural person. Based on that mentioned Law, even if several sorts of data compose digital assets, generally, when combined, they allow the identification of the individual holders, which characterizes the type of personal data. Thus, after exposing this definition, it is convenient to analyze data protection from the perspective of a person's right (defined by the Civil Code) and as a constitutional right. Person's rights are a recent legal category permeated with discussions about their existence, definition, and concepts. According to Limongi França[7], these rights can be defined as a law faculty whose object are the various aspects of the person, as well as their extensions; moreover, those rights aim to preserve human dignity. Some examples: the right to have a name, right to honor, right to image, right over one's own body, right to privacy, and many others. Bringing this theme closer to the central object of this work, the paragraph of article 20 of the Civil Code (CC) predicts that in the case of a deceased or absent person, the spouse, ascendants, or descendants are legitimate parties to request this protection. Therefore, although death ends the natural person, for private law, some remnants of personality endure and may be protected by the person's successors – it is known as *post-mortem* efficacy.

5 Diniz, L. (2015) 'O que são dados?', *Portal SlidePlayer*. Available at: https://slideplayer.com.br/slide/1745313/.

6 Ibid.

7 Szaniawski, E. (2005) 'Direitos de personalidade e sua tutela', *São Paulo: Editora Revista dos Tribunais*, 2 ed., p. 71.

In the essential conceptualizations, privacy is the "right to be let alone". A negative character endows this conception in the sense of preventing intervention in private life, just as it prevented intervention in a private property. From this analogy, it is possible to infer patrimonial content in the definition of privacy, which is strengthened by the vague notion that only a class with high purchasing power has privacy. This misconception, emerged from the initial conception of privacy, remains solid in nowadays mentality.

In opposition to that restricted view, Anderson Schreiber[8] affirms that this right must propose something more than that initial purpose which is restricted to the protection of intimate life. It should also refer to keep control over their personal data. Thus, it is understood that right to data protection is a person's right covered by the right to privacy which refers to the control of the gathering, filing, and use of personal data.

Recently, data protection has also become a constitutional right foreseen in the fifth article. So, it is essential to present the conception and the basic characteristics of constitutional rights in Brazil: firstly, constitutional rights are human rights inserted in a constitution of a specific State[9], which have limiting and justifying functions of State[10]. The limiting function prevents the State from overcoming such rights, while the justifying function is present in the purpose of State to implement them.

Furthermore, constitutional rights occupy a prominent role in the Constitutions, according to the classical thinking that goes back to the 18th century, since the two main objectives of constitutionalism would be: the limitation of power (with the organization of the State) and the achievement of constitutional rights and guarantees[11].

In the Brazilian Constitution, the 79th item was inserted in the fifth article through Constitutional Amendment no. 115 on February 10, 2022 (CA 115) and deals with right to protect personal data.

The GDPL previously mentioned already predicted the protection of personal data. Indeed, this is its fulcrum, but the inclusion of constitutional rights in the list gives the holder another mechanism to protect their rights according to Rosemberg Augusto Pereira Rodrigues, an analyst of the Federal Data Processing Service in Brazil. Therefore, a violation of data protection now constitutes a violation of a constitutional right, considering it is even stronger in a democratic context. Proceeding with the unfolding of the constitutional

8 Schreiber, A. (2013) 'Direitos da Personalidade', *São Paulo: Atlas*, 2 ed., pp. 135-136.

9 Sarlet, I. W. (2007) 'A Eficácia dos Direitos Fundamentais', *Porto Alegre: Livraria do Advogado Ed.*, 8 ed., pp. 35-36.

10 Sarlet, I. W. (2007) 'A Eficácia dos Direitos Fundamentais', *Porto Alegre: Livraria do Advogado Ed.*, 8 ed., p. 71.

11 Fernandes, B. G. (2021) 'Curso de Direito Constitucional', *Salvador: Ed. JusPodivm*, 13 ed., p. 39.

right to data protection for a digital inheritance, it is important to signalize some perspectives that approximate a person's rights and constitutional rights. In this regard, Kunrath[12] asserts that some jurists argue that it is not possible to consider a person's rights as constitutional once they are predicted in an ordinary law while constitutional rights are present in the Constitution. On the other hand, the author says that it prevails the understanding that a person's rights are constitutional rights if they are in accordance with the Constitution, since human dignity bases both types of rights[13]. Moreover, constitutional rights constitute principles that must be observed by all the norms of the same order. As the objective of this work is to analyze the digital inheritance in the light of the right to data protection – which is already established in the Constitution – it is not relevant to inquire deeper into the nature of constitutional rights and person's rights, being enough to know that the protection of data is considered in these two legal types. When it comes to digital inheritance, it is understood that there is a conflict between the constitutional right to inheritance – which grants the right to the heirs to receive the deceased's assets – and the constitutional right to data protection – which safeguards the privacy of the defunct. However, the aforementioned division linked to the patrimonial nature of digital assets would already resolve this impasse by legitimizing the transmission of assets with economic use and restricting the transfer of existential assets.

4. Responsibility for data protection

The responsibility for security in the virtual universe does not only belong to the State but also regards to companies, families and society. From this mutual responsibility it is evident that all social sectors need to develop ways to ensure the satisfactory use of technological tools. As this research deals with digital inheritance, the proposals developed here focused on protecting the deceased's data. Regarding the State, it is necessary to insert the processing of data of a deceased person into the legislation. Although the GPDL exists, it does not refer to this type of data, which creates insecurity about the proper destination, besides tarnishing the "general" character of law. About this omission, it is suggested to add a section in Chapter II – dedicated to the processing of personal data – about the Processing of Personal Data of Deceased Persons, such as there already is a special section for data related to children and teenagers.

Related to the constitutional right to data protection, the GPDL must predict that affective digital assets cannot be transmitted after death unless the latter

12 Kunrath, Y. C. (2016) 'Os direitos da personalidade enquanto direitos fundamentais', *Justiça do Direito, Passo Fundo*, vol. 30(3), pp. 503-522. Available at: http://seer.upf.br/index.php/rjd/article/view/6178/4019.

13 Ibid.

leaves a manifestation of will allowing the transmission. The privacy policies and terms of use cannot contain a provision for transmission to heirs of affective assets, if this occurs, the responsible person may be punished, as it is already defined in the GPDL. On the other hand, digital patrimonial assets are transmitted to the heirs, and it is a task of the person responsible for processing the company's data to correctly dispose them to the heirs.

Regarding the transmission of patrimonial assets, it is important to insert in Book V of CC - entitled On the Law of Inheritance, article 1.784 – the single transmission of digital heritage assets with economic expression. For example, based on the provision of article 1784, CC, it can be foreseen that once opened the succession, the digital patrimonial assets must be transmitted to the legitimate and testamentary heirs.

Ultimately, an addendum to Law No. 12,965 of 2014 (named in Portuguese "Marco Civil da Internet" – MCI) is suggested, which deals with the Internet use in Brazil. Currently, in its seventh article among the rights of users are the definitive exclusion of personal data when its owner requests after the end of relation. An exception is valid only when the files must be kept indefinitely or when the law obliges that. As personal data compounds the affective digital assets of deceased people – that is, data that needs an expression of will to be transmitted, otherwise, they must be excluded – it is important that this item clarifies the need to exclude this set of data from the *de cujus*, according to the cases of mandatory custody provided in MCI. It is also an obligation of the companies to adopt devices to protect personal data. Even without a legal provision for the transmission of digital assets of a deceased person, they can define protocols to organize the treatment of digital assets of this specific group, which distinguish them into affective and patrimonial assets, focused on the differentiated destination of each type, and defining forms to contact the deceased's family to discuss the transfer of assets. In addition, there is a new privacy protection paradigm that companies must adopt to avoid problems in this area, named "privacy by design" (PbD) also called "Privacy from Conception". The consideration of privacy protection during the entire useful life of a product, from the initial conception of the product until the end of its useful life, bases this model[14]. Thus, companies that adopt the PbD model develop privacy security mechanisms that follow all industrial and corporate activities. On the other hand, those companies that did not opt for this recent paradigm do not consider privacy protection important even in the design and production phase of the item, becoming extremely concerned when a problem arises, since it will possibly cause financial losses. Practices in line with the PbD model can be divided into administrative practices, such as staff training to respond to data

14 Bu, F.; Wang, N., Jiang, B., Liang, H. (2020) "Privacy by Design' implementation: Information system engineers' perspective', *International Journal of Information Management*, vol. 53. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0268401219308606

requests and investment in a privacy policy and terms of use of data that are clear and accessible to users; and technical practices (linked to the direct use of data), such as encryption, anonymization, and pseudo-anonymization[15]. Society and families, moreover, can work together in schools, social media and any other space conducive to discussions about the recent problem of the transmission of *post mortem* digital assets and care in the digital environment, including ways to improve the data protection. Finally, as there are legal instruments to control how data is used, the main user can choose the destination of their data, already preventing problems for the heirs with the filing of certain data. Among the measures provided by GPDL are: the correction of incomplete, inaccurate or outdated data (article 18, II) and the elimination of personal data processed with the consent of the holder (article 18, VI), except in the legal hypotheses of conservation for compliance with a legal obligation, study by a research institution, transfer to a third party and exclusive controller's use. Moreover, the user can express their desire regarding the data transfer or destination directly in the platform or application that contains it. Testament and codicil are also suitable legal documents to manifest his will regarding digital assets. There is also the possibility of using a secure digital will platform to guide how to file the data and to express the transmission.

For this research to constitute not only a theoretical orientation but also a practical one concerning the processing of the deceased's data, the suggestions previously presented are just non-exhaustive examples of measures that can be adopted by law, companies, social groups and individuals themselves to protect personal data.

5. Current status of digital inheritance: bills, court decisions and data policy

When analyzing the Law Draftings in progress in the Chamber of Deputies that aim to insert the digital inheritance in Brazilian law, it is possible to notice that there are many divergent positions between them. For example, the initiative n. 410/2021 aims, among other changes, to insert in the MCI the responsibility of internet application providers to delete the accounts of dead Brazilian users immediately after proof of death. This means that the project does not consider the possibility of transmitting digital patrimonial assets of the deceased without the existence of a manifestation of will.

The 3050/2020 initiative – which is being processed in commissions and has six attached Law Draftings – aims to amend article 1.788, CC, proposing the insertion of one paragraph in its text: "All contents of patrimonial quality, accounts or digital files owned by the author of the inheritance will be transmitted

15 Ibid.

to the heirs". It means that the transmission of digital assets must occur indiscriminately, ignoring the particularities of each case. It is relevant to highlight that those initiatives have predated CA 115, contributing to define the right to inheritance by the deceased's relatives. In this way, those initiatives in progress in the Legislative Houses need to be analyzed and discussed through the prism of data protection.

Regarding Courts decisions, there are examples that adopt the distinction of digital assets according to economic content. It is important to inform that this qualitative and quantitative research was executed in the State Courts of all Brazilian Southern and Southeastern States, and in the Court of Justice of the Federal District, using the terms "herança digital" (in english, "digital inheritance") and "transmissão de bens digitais" (in english, "transmission of digital assets") in the time from January 1, 2019 to August 30, 2022. As a result, it was possible to find only four judgments in the Sao Paulo State Court and one in Minas Gerais State Court, which was chosen due to fact that the explanation made about digital inheritance coincides with the differentiation of digital assets presented in this extended abstract, demonstrating that a uniform adoption of this view by the Courts is fully possible. The statement contains this fragment: "The inheritance refers to a unitary whole, which includes not only the material heritage of the deceased, but also the intangible, in which the digital assets of substantial economic valuation are inserted, called digital inheritance […]"[16]. Analyzing the content of the decision, it looks like judicial authorization is required to access private information, if relevant, which may indicate a phenomenon of the transformation of the nature of the digital asset, from affective to patrimonial, legitimizing the transmission. Metaverse, NFTs, cryptocurrencies and many other inventions of virtual reality show the complexity of dealing with the transmission of digital assets and that there are still many aspects to be defined about this issue.

Finally, it is essential that the law follows and adapts itself quickly with the new technologies, platforms and digital products. For this reason, the points previously exposed are proposed to encourage reflection and discussion.

16 Minas Gerais. Tribunal de Justiça do Estado de Minas Gerais. Agravo de Instrumento-Cv 1.0000.21.190675-5/001 1906763-06.2021.8.13.0000(1). Available at: https://www5.tjmg.jus.br/jurisprudencia/pesquisaPalavrasEspelhoAcordao.do?&numeroRegistro=1&totalLinhas=1&paginaNumero=1&linhasPorPagina=1&palavras=%2522heran%E7a%20digital%2522&pesquisarPor=ementa&orderByData=2&referenciaLegislativa=Clique%20na%20lupa%20para%20pesquisar%20as%20refer%EAncias%20cadastradas...&pesquisaPalavras=Pesquisar&.

6. Conclusion

The analysis presented in this work – the digital inheritance considering the constitutional right to data protection – aims to indicate a way to solve the succession impasse about the interests of the heirs and, at the same time, defends the privacy of the author in front of the inheritance of the digital assets.

The importance of dividing digital assets into patrimonial and affective assets is highlighted to indicate the correct destination of data with economic value, while affective assets require a manifestation of will.

In order to properly treat all the data of a deceased person, it is necessary a legislative change, besides the act to improve data protection mechanisms and facilitate transmission, with a manifestation of will, for example, as a society.

This way solves the existing omission, adapts the situation to the constitutional right to data protection, and contributes to the development of digital law, that corresponds to the scope of defining responsibilities in virtual relationships.

The fact is: death is inexorable, but most problems that can arise with it, especially regarding to digital assets, can and must be avoided, for example, by using the mechanisms mentioned in this extended abstract.

PART III
Big Data, Platforms Regulations and Open Data

Chapter XII
Legal and Ethical Challenges in the Use of Web 2.0 Open Data

by Jonida Milaj*

INDEX: 1. Introduction. – 2. Challenges in the use of Web 2.0 Open Data. – 3. The Privacy and the Data Protection framework. – 4. Web 2.0 Open Data and AI. – 5. Conclusion.

1. Introduction[1]

The use of open-source data from social media, focuses on extracting insights from publicly available data in Web 2.0[2] platforms with a focus on micro-blogging (Twitter), video-sharing (YouTube), social-networking (Facebook), etc. In the European Union, the use of open data from the public sector for training Artificial Intelligence (AI) is currently encouraged and supported in the framework of the Digital Single Market Strategy[3]. Closely linked to this, the Data Governance Act[4] aims to further facilitate the reuse of public sector data as well

---

* Assistant Professor in Technology Law and Human Rights at the University of Groningen (the Netherlands) and deputy academic director of the LLM programme in Technology Law and Innovation. Her main research is hosted by the the Security, Technology and e-Privacy (STeP) research group and it focuses on the challenges that data driven innovation and technology developments create for the protection of fundamental rights of individuals. Jonida has widely published in renowned peer reviewed international journals and edited volumes. She is a research fellow at the Information Society Law Center of the University of Milan (Italy) and a visiting lecturer at the Central University of Political Science and Law in Beijing (China).

1 Part of this research has been conducted in the framework of H2020 CRiTERIA project (Comprehensive data-driven Risk and Threat Assessment Methods for the Early and Reliable Identification, Validation and Analysis of migration-related risks - Grant Agreement № 101021866).

2 The terms 'social media' and 'Web 2.0' are used interchangeably in this Paper.

3 Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A digital Single Market Strategy for Europe [2015] COM/2015/0192 final.

4 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) [2020] OJ L152/1.

as to facilitate and encourage data sharing for individuals as well as for private businesses, based on the data altruism concept[5].

In this framework, the use of open-source Web 2.0 data is also encouraged[6]. The use of these data is especially crucial for training algorithms and developing new AI solutions.

There are no specific rules that regulate the use of open data from Web 2.0. Thus, their use falls under the general legal framework on data protection. While the General Data Protection Regulation (GDPR)[7] introduces safeguards for processing of personal data, these safeguards are watered down for Web 2.0 open-source data based on two assumptions: (i) individuals that publish their data without any privacy filters do not have any reasonable privacy expectation; (ii) there are no legal restrictions on processing personal data that are made manifestly public by data subjects themselves. Thus, despite the "trusted" data-sharing tools that the legislator has introduced, the real protection of the rights of individuals is at a crossroad.

By challenging the mainstream view, this research takes a fundamental right, as well as a multi-disciplinary approach to address legal and ethical challenges for the use of opensource data from social media. The use of these data, without adequate safeguards, brings the downfall of the human rights protection system as we know it. Thus, the interference with the rights of individuals when processing Web 2.0 open data goes far beyond privacy and data protection concerns. It affects their freedom of expression, association and religion, it undermines the fair and due process, it affects the autonomy and dignity of data subjects, it creates chilling effects in society and thus, it undermines democracy.

After this introduction, section 2 will focus on the legal and ethical challenges identified in the use of Web 2.0 open data. In section 3, the data protection and privacy regulation of open data will be analysed. In section 4, the draft AI Act is discussed with regards to any insights relevant from an open data perspective. Section 5 will conclude reflecting upon the challenges that the use of Web 2.0 open data creates for the system of human rights protection. Potential solutions and safeguards in light of the European Data Strategy[8] are suggested.

---

5 Data Governance Act, art 2(16).

6 The European Space Agency wrote that: "[…] analysing social media data allows for better understanding of the behaviour and sentiments of crowds at a particular geographic location and a specific moment in time, which can be indicators of possible migration movements in the immediate future".

7 Regulation (EU) 2016/679 0f 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 the free movement of such data, and repealing Directive 95/46/EC (GDPR) [2016] OJ L119/1.

8 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 [2020] COM/2020/0066 final.

2. Challenges in the use of Web 2.0 open data

Open-source data from social media are seen as providing useful sources for training algorithms. However, with the advantages come also many challenges. These challenges might be of a legal nature, but not only. When responsibly addressing new developments in technology that are related with data, legal compliance is just one part of a much bigger picture[9]. Further ethical concerns come into play. Ethics goes far beyond the laws to consider ways of behaviour that would not cause harm to others, especially in the presence of legal lacunas[10].

The publication of certain illegal uses of online data, as for example the Facebook's experiments in emotion manipulation[11], or the use of social media by data analytics companies seeking insights into citizens' political attitudes and networks to influence voter behaviour[12], have raised awareness on the potential of these data, but there is no evidence as to what extent they have influenced individuals' online behaviour.

When processing data from social media, the complexity of interactions between individuals, groups and technical systems becomes very relevant[13]. From the perspective of data input, these complexities include: i) the self-selecting nature of social media users; and ii) inequalities in access to social media platforms and data. From the processing of data and an output perspective, complexities are identified with regards to: iii) the difficulty to obtain meaning from heterogeneous data of variable quality and provenance; iv) potential limitations of systems in the amount of data provided; and a v) dependency on observing and interpreting what is "out there" in a way that differs from traditional approaches[14].

Currently, many challenges that arise when open data from social media are used, have an ethical nature[15]. For social media users, for example, there might

9 Hijmans, H., Raab, C. (2018) 'Ethical Dimensions of the GDPR' in: Cole, M., Boehm, F. (eds), *Commentary on the General Data Protection Regulation*. Edward Elgar. Available at: https://ssrn.com/abstract=3222677.

10 French, C. (1893) 'The Concept of Law in Ethics', *The Philosophical Review*, vol. 2(1).

11 Jouhki, J. et al. (2016) 'Facebook's Emotional Contagion Experiment as a Challenge to Research Ethics', *Media and Communication*, vol. 4(4), pp. 75-85; Jonida Milaj, J., Bonnici, J.M. (2022) 'Stitching lacunas in open-source intelligence – Using Ethics to fill up legal gaps', *Illyrius*, vol.18(1), pp. 47-57.

12 Isaak, J., Hanna, M.J. (2018) 'User Data Privacy: Facebook, Cambridge Analytica, and Privacy Protection', *Computer,* vol. 51(8), pp. 56-59.

13 Munson, S.A. et al. (2013) 'Sociotechnical challenges and progress in using social media for health', *J. Med Internet Res*, vol. 15(10), e226.

14 Taylor, J., Pagliari C. (2018) 'Mining social media data: How are research sponsors and researchers addressing the ethical challenges', Research Ethics, vol. 14(29), pp. 1-39.

15 Berry, D.M. (2004) 'Internet research: privacy, ethics and alienation: an open-source approach', *Internet Research*, vol. 14(4), pp. 323–332.

be variable perceptions and unclear boundaries between "public" and "private" spaces. They might consider their online presence as private, being connected only to a limited number of "friends" or "followers", while they have failed to introduce the proper privacy settings in their accounts. As a result, their social media activity qualifies as open source. Furthermore, processing open data from social media has impediments in ensuring full anonymity and in preserving the privacy of data subjects. Identities may be difficult to disguised or may be easily deduced from public postings and affiliations. Social media data might also reveal directly or indirectly[16] sensitive information as well as data from vulnerable data subjects[17]. When processed, these data would require more stringent safeguards than the rest of the massive amount of open-source data.

All the above-mentioned challenges derive from the nature of social media. The risk is present that, depending on the purposes of processing, or the algorithms that will be trained, these challenges translate into further human rights and ethical concerns for individuals. Firstly, social media users are given a digital identity. From the algorithm training point of view, users are relevant with regards to their online behaviour at specific points in time. These are the moments when data are collected and processed. The technical time fragmentation of the collection and processing of data creates potential risks and implications for the protection of the social media users autonomy and dignity. We are all aware that Web 2.0 allows for the creation of various profiles and identities, depending on the friends/followers base[18]. These sociological nuances cannot be reflected in the data harvested and processed. Secondly, the inherent challenges that derive from the nature of open data from social media, might potentially create situations of discrimination or biased data processing. Thirdly, the awareness of the use of social media data for training algorithms, for example with regards to risk assessments, might have as a consequence that individuals use social media less, are not expressing their opinion freely or are using coded language. Thus, chilling effects are created in the society. Fourthly, processing Web 2.0 data might bring the creation of stereotypes. This risk might be augmented by the fact that users of social media platforms are self-selected, and inequalities exist in accessing the services. The bias effect might be even the result of the way data are received by social media, thus might be inherent of another independent existing system. Last but not least, users might change their mind or previously posted information. Moreover, in order to understand and qualify information, there might be the need to understand the context

16 Case C-184/20 Vyriausioji tarnybinės etikos komisija EU:C: 2022:601, para 128.

17 Kuyumdzhieva, A. (2018) 'Data Ethics and Ethics Review Process. Ethics compliance under GDPR', *Bioethica*, vol. 5(1). Presentation available online: https://ec.europa.eu/info/sites/default/files/6._h2020_ethics-soc-science-humanities_en.pdf.

18 Katz, R., Ogilvie, S., Shaw, J., Woodhead, L. (2021) 'Gen Z, Explained: The art of living in a digital age', *The University of Chicago Press*.

and mental framework in which certain data were published. This cannot be ensured for massive data collection and processing activities. Thus, the risk of using untrustworthy data and databases is present.

3. Privacy and Data Protection considerations

The legal framework within which the use of open data from social media for training AI takes place is broad and, it includes copyright, tort law, product liability, as well as fundamental rights, especially with regards to the rights to privacy and data protection. Given the lack of space to address every and each of the relevant legal fields, this section will address the rights to privacy and to data protection. Especially, the two assumptions at the basis of processing open data from social media: i) individuals that publish their data without any privacy filters do not have any reasonable privacy expectations; and ii) there are no legal restrictions on processing personal data that are made manifestly public by data subjects themselves; will be addressed.

The right to Privacy and Web 2.0 open data

The right to a protected private life is defined in article 8 of the European Convention on Human Rights (ECHR) and article 7 of the Charter of Fundamental Rights of the European Union (CFR). States do not only have a negative obligation, not to interfere with the private sphere of individuals, but also a positive one, to ensure that there are laws in place that safeguard from any arbitrary interference[19]. Allowing or suggesting and supporting the use of open-source data from social media for training AI would fall under the positive State obligations. This sub-section focuses on the right to privacy with regards to open-source data and on the existence or not of a reasonable expectation of privacy when open data are processed.

Clearly, the right to privacy does not protect individual only behind closed doors. In Perry the European Court of Human Rights confirmed that the right to privacy exists also outside a person's home or private premises[20]. In P.G. and J.H., the Court established that private-life considerations may arise for activities in physical open spaces once a systematic or permanent record of data comes into existence[21]. Although the above judgments were addressing the physical world, there are no reasons why the same logic should not apply to cyberspace[22]. However, the systematic collection or permanent storage qualifies

19 Marckx v. Belgium App no 6833/74 (ECtHR, 13 June 1979), para 31.

20 Perry v The United Kingdom App no 63673/00 (ECtHR, 17 July 2003), para 37.

21 PG and JH v. The United Kingdom App no 44787/98 (ECtHR, 25 September 2001), para 57.

22 Edwards, L., Urquhart, L. (2016) 'Privacy in public spaces: What expectations of privacy do we have in social media intelligence?', *International Journal of Law and Information Technology*, vol. 24, p. 279.

the processing of open data as an interference with the private sphere of individuals and doubts remain with regard to non-systematic data collection and processing[23].

The "reasonable expectation of privacy" principle has its origins in the United States and the British case law[24]. In Katz, the US Supreme Court introduced a two steps test including: (i) a subjective expectation of privacy in certain situations, and (ii) an objective expectation linked to the recognition of the expectation from the society[25]. The European Court of Human Rights has used the principle as one of the ways for establishing if an infringement of the right to privacy exists. In Perry, the Court reasoned that an individual has a reasonable expectation of privacy if he is not being able to reasonably expect the use of technology for scopes beyond the normal foreseeability of its use[26]. We argue the same reasoning applies also to the processing of open data from social media. Social media users do not expect the use of their data and technology beyond their original scope of communication and thus, the reasonable expectation of privacy principle would be fulfilled.

However, the right to privacy is not an absolute one[27]. The interference with the private sphere of individuals could be justified by lawful grounds, as for example when training AI is done in the public interest or for the economic well-being of the country. However, the challenges and risks of processing open data from social media that were identified in section 2 of this paper, would raise serious doubts about the necessity and proportionality of the interference. A case-by-case assessment needs to be conducted.

The right to Data Protection with regards to Web 2.0 open data

Open-source data from social media might fall under the category of personal data[28]. Because it is relatively easy to go back to an individual post, anonymization would not change this qualification for as long as there is the possibility to create a mosaic effect and to identify a data subject by connecting different

23 Koops, B.-J. (2013) 'Police investigation in Internet open sources: procedural-law issues', *Computer Law & Security Review*, vol. 29(6), p. 654.

24 Gomez-Arostegui, T. (2005) 'Defining Private Life Under the European Convention on Human Rights by Referring to Reasonable Expectations', *California Western International Law Journal,* vol. 35(2), pp. 153-202.

25 Katz v. United States 389 US 347 (1967).

26 Perry v. The United Kingdom App no 63737/00 (ECtHR, 17 July 2003), para 41.

27 Kleining, J. et al. (2011) 'Security and Privacy: Global standards for ethical identity management in contemporary liberal democratic states9, *ANU E Press*, vol. 43; Kilkelly, U. (2003) *The right to respect for private and family life,* p.6. Available at: http://echr.coe.int/NR/rdonlyres/77A6BD48-CD95-4CFF-BAB4-ECB974C5BD15/0/DG2ENHRHAND012003.pdf; Kenneth, E., Himma, K.E, (2007) 'Privacy vs Security: Why privacy is not an absolute value or right', *San Diego Law Review*, vol. 44, p. 857.

28 GDPR, article 4(1).

databases. As a result, the general data protection framework established by the GDPR is applicable.

Open-source data are not explicitly addressed in the GDPR, thus they fall within the general legal framework. For complying with the data protection framework, first attention must be paid to the principles of lawful data processing established in article 5 GDPR, namely: (a) lawfulness, fairness and transparency; (b) purpose limitation; (c) data minimization; (d) accuracy; (e) storage limitation; (f) integrity and confidentiality; and (g) accountability. The use of open data from social media might raise direct concerns with regards to the principles of transparency, purpose limitation, as well as data minimization and accuracy. While the potential problems with regards to transparency, data minimization and accuracy are directly linked to the challenges of use of open-source data, the discussion of the purpose limitation principle is more technical. It derives from the nature of social media and their use by individuals for communication purposes – data made available for the purpose of communication are now used for training AI. However, to limit the use of these data only to communication purposes would have the extreme result that these available data can never be used for other purposes as for example for research, unless there is the explicit consent of data subjects to use these data for research purposes and would contradict the current EU Digital Market Strategy. Article 5(1)(b) GDPR clarifies that the use of personal data in the public interest is not considered to be incompatible with the initial purposes of data processing.

Also, the assumption that the data are made manifestly public, thus any processing is in compliance with the GDPR, is not easily granted. Firstly, the blurring borders between public and private space on social media are difficult to be identified. This makes it difficult to establish if a data subject has willfully made the data available[29]. Secondly, the condition of consent cannot be established. Making personal data available does not automatically qualify as giving the consent to whomever has access to these data to process them for purposes different from the original intention of the data subject. Thirdly, some data might provide sensitive information[30]. A personal image, for example, might reveal a health condition, the religion of the data subject or his/her ethnic origin. Although the processing activity might not aim for revealing sensitive information, these might be an indirect result[31]. Processing can be justified on the basis of article 9(2)(e) GDPR on data that are made manifestly public. However, this needs to be considered in a restrictive way and always in combination with the fulfilment of the conditions for lawful data processing[32].

29 Case T-320/02 *Esch Leonhardt v ECB* EU:T:2004:45.

30 GDPR, art 9.

31 Case C-184/20 *Vyriausioji tarnybinės etikos komisija* EU:C:2022:601, para 128.

32 Edwards, L., Urquhart, L. (2016) 'Privacy in Public Spaces: What Expectations of Privacy Do We Have in Social Media Intelligence?', I*nternational Journal of Law and Information Technology*,

From the above analyses, it can be concluded that one needs to conduct a case-by-case analyses to ensure that the use of Web 2.0 open data is done in compliance with the rights to privacy and to data protection. The two assumptions on lack of reasonable privacy expectation and data made manifestly public also require a case-by-case evaluation. However, the large amount of Web 2.0 open data and the uncertainties linked to these data, make such assessment difficult, or even currently impossible.

4. Web 2.0 Open Data and AI

After analysing the privacy and data protection legal framework, it is important to analyse the way Web 2.0 open data are considered with regards to the AI rules. In April 2021[33], the draft AI Act was proposed by the European Commission[34]. With the future introduction of this Act, the EU aims to address the risks generated by specific uses of AI through a set of complementary, proportionate and flexible rules, and to have a leading role in setting the global standards for AI. The Act follows a risk-based approach, differentiating between uses of AI that would create for individuals:

- an unacceptable risk;
- a high risk
- a low or minimal risk.

While AI presenting unacceptable risks must not be used, high-risk AI is allowed while following some strict safeguards before being put on the market and also once the system is available on the market. For high-risk AI systems, the requirements of high-quality data[35], as well as documentation and traceability, transparency, human oversight, accuracy and robustness, are necessary to mitigate the risks to fundamental rights and safety posed by AI. Especially if these are not covered by other existing legal frameworks.

The high-quality data requirement is explained further in Recital 38, stating that if the AI system is not trained with high quality data, it does not meet adequate requirements in terms of its accuracy or robustness, and thus, it may single out people in a discriminatory or otherwise incorrect or unjust manner. Furthermore, the exercise of important procedural fundamental rights, such as

vol. 24, p. 279.

33 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 [2021] COM/2021/206 final.

34 At the time of writing (February 2024), the official draft of the AI Act available does not reflect on the trialogue agreement reached between the Commission, The European Parliament and the Council on the 9th of December 2023. The main outcomes of the political deal can be found here: https://www.europarl.europa.eu/news/en/press-room/20231206IPR15699/artificial-intelligence-act-deal-on-comprehensive-rules-for-trustworthy-ai.

35 Draft AI Act, art 10 and rec 45.

the right to an effective remedy and to a fair trial as well as the right of defence and the presumption of innocence, could be hampered, in particular, where such AI systems are not sufficiently transparent, explainable and documented. This is reinforced further in Recital 44 where it is stated that high-quality data is essential for the performance of many AI systems, especially when techniques involving the training of models are used, with a view to ensure that the high-risk AI system performs as intended and safely, and it does not become the source of discrimination prohibited by EU law.

In section 2 of this paper, it was argued that the use of open data from social media suffers from a number of inherent challenges that are reflected in the quality of the results of processing these data. Thus, Web 2.0 open-source data cannot fall under the high-quality data category and cannot be used in training AI systems that presents high risks and directly affect the rights of data subjects. Furthermore, the identified inherent problems of social media data also mean that using these data for training minimal or low risk AI would have the consequence of potentially discriminating individuals and thus they should also not be used in these cases as to avoid any potential discrimination or harm[36].

Web 2.0 open data will create problems also with regards to ensuring the transparency of AI decisions. Transparency is focal in the draft AI Act. Furthermore, in the judgment of the CJEU with regards to compatibility of the Passenger Name Record (PNR) Directive[37] with fundamental rights[38], the Court specified that automated systems used to identify suspicious individuals must be based on objective and non-discriminatory criteria. A human review is needed to verify the flagged individuals. Automated systems cannot use machine-learning techniques capable of modifying, without human intervention or review, the assessment process and, in particular, the assessment criteria on which the result of the application of that process are based as well as the weighting of those criteria, because "given the opacity which characterises the way in which artificial intelligence technology works, it might be impossible to understand the reason why a given program arrived at a positive match"[39]. In light of this judgement, as well as of the inherent challenges in Web 2.0 open data already discussed, the use of these data for training algorithms does not allow for ensuring the transparency of the processing operations.

36 Bostrom, N. (2003), *Ethical Issues in Advanced Artificial Intelligence*, in Smit et al. (eds) *Cognitive, Emotive and Ethical Aspects of Decision Making in Humans and in Artificial Intelligence,* International Institute of Advanced Studies in Systems Research and Cybernetics, pp. 12-17.

37 Directive (EU) 2016/681 of the European Parliament and of the Council of 27 April 2016 on the use of passenger name record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime [2016] OJ L119/132.

38 Case C-817/19 Ligue des droits humains EU:C:2022:491.

39 Case C-817/19 Ligue des droits humains EU:C:2022:491, paras 194-195.

5. Conclusion

Despite their large amount, availability and the easy accessibility, the use of Web 2.0 open data for training AI systems needs to be handled with care since it presents a number of legal and ethical challenges. The use of these data creates problems for the protection of the autonomy and dignity of social media users, might create situations of discrimination or biased data processing, might create chilling effects for the society and, last but not least, might result in the creation of untrustworthy databases.

As it was argued in this paper, the assumption of lack of any reasonable privacy expectation for Web 2.0 open data is not legally supported. For compliance with the right to privacy, a case-by-case approach is needed to ensure that any interference with the private sphere of individuals follows the principles of necessity and proportionality. Furthermore, also the assumption of lawful processing since the data are made manifestly public by data subjects needs to be assessed following a case-by-case approach. Thus, potential legal problems can be identified due to lack of compliance with the data protection principles.

The application of the privacy and data protection standards is not affected by the future entry into force of the AI Act. These legal frameworks will continue to apply side by side. A legal basis under article 6 GDPR for the processing of personal data for training AI will be needed as well as compliance with the data protection principles listed in article 5 GDPR. In this paper it was argued that given all their challenges and legal limitations, Web 2.0 open data do not comply with the requirement for high-quality datasets highlighted in the new AI legal framework.

As a result of these concerns, despite their availability and current encouragement in using Web 2.0 open data for training AI, these data should not be used for training high-risk systems that directly affect the protection of fundamental rights of data subjects, unless there is a proper legal recognition and technical addressing of the legal and ethical challenges that the data present. Even though the current data protection and privacy legal framework applies to Web2.0 open data, a watering down of the legal safeguards is present. To avoid this weakened legal protection, a special law regulating the processing of data from social media in general and of open-source data in particular is needed. This law needs to specify the instances in which the data can be used based on specific legal and technical impact assessment analyses. This law will be crucial not only for protecting the rights to privacy and data protection of individuals while AI technology is developing with tremendous speed, but it will also ensure the protection of other fundamental rights on which our societies are based upon, such as: freedom of expression, information, association and religion, the fair and due legal process, the autonomy and dignity of data subjects and, above all, democracy.

Chapter XIII
How Smart Cities Leverage the Power of Data and Sensors to Bridge Digital Gaps and Foster Prosperity

by Beatrice Bonami*

INDEX: 1. Introduction. – 2. The digital age of smart systems. – 3. Globalization, data, sensors and citizenship. – 4. Digital rights. – 5. Final considerations: building a sustainable future with smart cities.

1. Introduction

The idea that Smart Cities can leverage the power of big data and sensors to bridge digital inclusion gaps, foster vulnerable youth groups to success and bring prosperity in the future economy is not far from reality. When discussing smart cities, most people envision a futuristic metropolis full of innovative technologies, such as intelligent traffic lights and self-driving cars. Science Fiction movies and literature (such as Blade Runner[1] or Her[2]) have played an essential role in how we imagine our interaction with digital technologies in the future – bringing humans closer to devices and merging their lives with artificial intelligent forms. While these technologies are undoubtedly helpful for improving the way we live and work, there is another aspect of smart cities that is often neglected – namely, the role they play in fostering youth engagement and digital literacy. Cities worldwide face enormous challenges in ensuring that

* PRIME Lead Researcher - Universität Tübingen/DAAD – bonamibeatrice@gmail.com. Dr. Beatrice Bonami is an Italo-Brazilian author, social researcher, and innovator. She has extensive multi-country experience in various multicultural settings, including government, educational environments, and indigenous territories. She holds a Ph.D. in the field of Innovation and International Development by the University of São Paulo [Brazil], the University College London [United Kingdom] and the Università La Sapienza di Roma [Italy]. She is a PRIME Lead Researcher at Universität Tübingen, holding a project about Intersectional AI in Africa.

1 *Blade Runner* (1982) science fiction film directed by Ridley Scott and written by Hampton Fancher and David Peoples. Starring Harrison Ford, Rutger Hauer, Sean Young, and Edward James Olmos, it is an adaptation of Philip K. Dick's 1968 novel Do Androids Dream of Electric Sheep?

2 *Her* (2013), directed by Spike Jonze. With Joaquin Phoenix, Lynn Adrianna Freedman, Lisa Renee Pitts, Gabe Gomez.

their population has access to technology services, from being able to read and write to having access to modern forms of communication. Especially in emerging economies, many people live far away from urban centers and cannot access reliable broadband infrastructure or computing facilities[3]. This creates immense challenges when it comes to offering educational opportunities that meet people where they are and providing access to health care and other services that require connectivity. On the other hand, intelligent cities and neighborhoods are taking on the challenge of increasing access to technology by deploying a combination of data, sensing, and networking technologies that can foster greater inclusivity and promote social engagement among youth who are still left behind by the digital revolution. As the world becomes increasingly digitized and citizens have access to an ever-growing number of digital services and devices, we must protect their rights and ensure they can exercise their rights online[4]. But what exactly do we mean by digital rights, and what are these issues' significance for our future? And how are Smart Cities and the policies around sensors related to that? With this scenario and questions in mind, this conference paper analyzes examples of smart cities in Asia, Latin America, and Europe that have combined sensors and Big Data to create new learning spaces and opportunities for youth beyond the classroom and the community.

2. The digital age of smart systems

Smart cities are populational complexes that use technology to improve the living conditions of their citizens. By gathering data from connected devices in cities and implementing it in intelligent systems and services, cities can promote sustainability and sustainable development and provide better services to their residents. For example, a smart city could use the data of its citizens to determine where there are traffic jams on the roads and then implement traffic control measures accordingly to reduce traffic congestion[5].

It also makes use of data to monitor electricity and water consumption in public buildings and inform the public about the best possible ways to reduce consumption. These services include tracking passengers on public transport routes and providing them with information like wait times for the next bus or train[6]. Also, some houses are integrated with apps and smart devices that connect personal data to house functioning to offer control, efficiency, and

3 Ratti, C., Claudel, M. (2016) 'The City of Tomorrow: Sensors, Networks, Hackers, and the Future of Urban Life', *Yale University Press: United States*.

4 Ibid.

5 Jain, A. (2019) 'Smart Cities: From Vision to Action', *Discovery Publishing House: United States*.

6 See https://impact.economist.com/smartercities/?utm_medium=cpc.adword.pd&utm_source=google&ppccampaignID=18151738051&ppcadID=&utm_campaign=a.22brand_pmax&utm_content=conversion.direct-response.anonymous&gclid=Cj0KCQiAz9ieBh-

an improved user experience[7]. These smart technologies are already used in several cities like Singapore, Hamburg and Barcelona. The development of smart cities raises a number of challenges, such as the security of user data and privacy issues as well as ethical concerns about the use of information and data collected in the city[8].

The concept of smart cities is relatively new, and there is no clear consensus on which city can be considered the first smart city However, several cities have been recognized as early adopters and leaders in the development of smart city technologies. The concept of smart cities can be traced back to the late 1990s and early 2000s. The term smart city started to be used in reference to the use of technology to improve the management and operation of cities and to improve the quality of life of citizens. One of the earliest references to the concept of a smart city can be found in a 1999 article by Kevin Ashton, the Executive Director of the Auto-ID Center at MIT[9]. In the article, Ashton proposed the use of radio frequency identification (RFID) technology to create an Internet of Things that would connect everyday objects and make them smarticle He argued that this technology could improve cities' management and operation and citizens' quality of life. Another early reference to the concept of a smart city is from a 2000 report by the European Union, which defined a smart city as a city that excels in the use of information and communication technologies (ICT) to improve quality of life, the efficiency of urban services and competitiveness, while ensuring that it meets the needs of present and future generations concerning economic, social, and environmental aspects. Since then, the term smart city has been widely adopted and used to refer to the use of technology to improve the management and operation of cities and to improve the quality of life of citizens[10]. It's important to note that the concept of smart cities has evolved, and different people and organizations may have different definitions and interpretations of the term. In a similar direction, in 2023, it is already possible to note interesting examples of smart cities and neighborhoods. One example is Songdo[11], South Korea, which was developed

---

CIARIsACB0oGJJ4D52FfEx_YOjk6KrcnKm1za2IaXpOlv-m7HUNuHna66WLNrLABEaAslpEALw_wcB&gclsrc=aw.ds.

7 Quitzow, L. (2022) 'Smart grids, smart households, smart neighborhoods – contested narratives of prosumage and decentralization in Berlin's urban Energiewende', *Innovation, The European Journal of Social Science Research*. Available at: https://doi.org/10.1080/13511610.2022.2057934934.

8 Jain, A. (2019) 'Smart Cities: From Vision to Action', *Discovery Publishing House: United States*.

9 Ashton, K. (2009) 'That 'Internet of Things' thing', *RFID J*. Available at: http://www.rfidjournal.com/articles/view?4986.

10 Kitchin, R., Dodge, M. (2012) 'Code/Space: Software and Everyday Life (Software Studies)', *MIT Press: United States*.

11 See https://www.archdaily.com/962924/building-a-city-from-scratch-the-story-of-songdo-korea.

as a smart city from the ground up. The city was built on reclaimed land near Incheon, South Korea, and was designed to be a model of sustainable, high-tech urban living[12]. The city incorporates technology and data into every aspect of urban life, from transportation and energy systems to buildings and public spaces. The South Korean government developed the city and a consortium of private companies, led by Gale International and POSCO E&C, to create a city that could serve as a model for sustainable urban development in the 21st century. Songdo is a pioneer and an example of smart city development. Still, it faced challenges, such as the high cost of living, lack of community and culture, and the need to attract more residents and businesses to the city. The city's infrastructure incorporates cutting-edge technologies such as high-speed wireless networks, advanced building automation systems, and a city-wide transportation management system. The concept of a smart city is closely tied to the development of digital technologies and the internet, which is most commonly associated with the digital age. However, the idea of creating more efficient and livable cities through technology and data has existed for much longer. One example is the Garden City movement[13], which originated in the late 19th and early 20th centuries. The movement sought to create new towns and cities that would be planned and built around the principles of urban planning and garden design, intending to provide a healthier and more sustainable way of life for residents. The Garden City movement emphasized the importance of green spaces, efficient transportation systems, and community-oriented design in creating livable cities. Another example is the concept of the intelligent city that emerged in the late 20th century, which refers to cities that can use advanced technologies and data to improve residents' efficiency and quality of life. Intelligent cities typically focus on transportation, energy, and public services and aim to provide citizens with more efficient and convenient services. In summary, the current smart cities are closely tied to the digital age; the concept of creating more "livable" and efficient cities through the use of technology and data has existed for much longer and has evolved.

In terms of role models, Singapore is a leading smart city. The city-state has implemented several smart city initiatives that aim to improve the quality of life for residents and visitors and make the city more sustainable and efficient. Some examples of Singapore's smart city initiatives include smart mobility, a comprehensive public transportation system that includes buses, trains, and ferries, bike-sharing, and autonomous vehicle technologies.

---

12 It was created by a Public-Private Partnership (PPP) between the Incheon Free Economic Zone Authority and private partners, Gale International and POSCO E&C, with over $35 billion in investment. The development of the city began in 2002 and was completed in 2014. It is now home to approximately 40,000 residents.

13 See https://www.britannica.com/topic/garden-city-urban-planning.

Singapore is committed to protecting the environment and reducing its carbon footprint. The city has implemented measures such as the use of renewable energy sources, the promotion of sustainable transportation, and the conservation of green spaces. The city has also implemented several digital services and platforms that improve the transparency and efficiency of city government and a highly developed and advanced security system, which is supported by technology and data analytics to enhance the safety and security of its citizens.

Singapore is often cited as a model for other cities to emulate and is widely recognized as one of the most advanced smart cities in the world. The government is actively investing in technology and infrastructure to improve its residents' and visitors' quality of life. Today's smart cities are no longer defined by new technologies or particular aspects of urban planning and development. Instead, they are models in which technology enhances efficiency and improves sustainability across various activities – from city planning to citizen engagement to service provision.

3. Globalization, data, sensors and citizenship

Other continents aside from Asia have been investing in creating intelligent forms of habiting the 21st Century. Several cities in Latin America and Europe, for example, have been recognized as leaders in developing smart city technologies. Some examples include:

1. Medellin, Colombia: Medellin is often cited as a model for smart urban development in Latin America. The city has implemented several innovative urban development projects, such as constructing a cable car system to connect disadvantaged neighborhoods to the city center and creating some public spaces, such as the Botanical Garden, which has become a major tourist attraction.
2. Santander Smart City, Spain: Santander is a pioneer city in the development of smart services for citizens, such as the "Santander Smart Bikes", the "Santander Smart Parking", the "Santander Smart Lighting" and the "Santander Smart Traffic". It also has a citizen engagement platform called "Santander Participa", which allows citizens to propose and vote on new services and urban improvements.
3. Quito, Ecuador: Quito has implemented many smart city initiatives, such as the installation of a network of environmental sensors to monitor air and water quality, and the creation of a mobile application that allows residents to report and track issues such as potholes and broken streetlights. The city also has a comprehensive transportation management system that allows residents to plan their routes, view traffic conditions and pay for parking.
4. Montevideo, Uruguay: Montevideo is considered a pioneer in the use of technology to improve public services, with initiatives such as the

"Montevideo Digital" program, which aims to provide all citizens with access to high-speed internet and digital services. The city also has a platform called "Montevideo Participa", which allows citizens to propose and vote on urban projects, and a "Smart Lighting" system, which aims to optimize energy consumption and improve safety in public spaces.

5. Buenos Aires, Argentina: Buenos Aires has been working on a number of smart city initiatives, such as the implementation of a comprehensive transportation management system, which includes a network of cameras to monitor traffic conditions and a mobile application that allows residents to plan their routes, view traffic conditions and pay for parking. The city also has a smart lighting system that adjusts the brightness of streetlights based on the amount of ambient light.
6. Amsterdam, Netherlands: Amsterdam is known for its innovative approach to urban planning and its use of technology to improve the quality of life for residents. The city has implemented a number of smart city initiatives, such as the Amsterdam Smart City platform, which uses data and analytics to improve city services.
7. Barcelona, Spain: Barcelona is considered to be one of the leading smart cities in Europe, and has been recognized for its smart city initiatives, such as the creation of a city-wide Wi-Fi network and the development of a smart grid for energy management.
8. Copenhagen, Denmark: Copenhagen is considered one of the world's most sustainable cities and is known for its smart city initiatives, such as its bike-sharing program, which uses technology to improve the efficiency and accessibility of the program.
9. London, United Kingdom: London is one of the world's most innovative and vibrant cities. The city has implemented a number of smart city initiatives, such as the creation of a city-wide Wi-Fi network, the development of a smart grid for energy management, and the use of data and analytics to improve city services.
10. Vienna, Austria: Vienna is known for its smart city initiatives, such as the development of a smart grid for energy management, the use of data and analytics to improve city services, and the implementation of a city-wide bike-sharing program.
11. Hamburg, Germany: Hamburg is considered to be one of the leading smart cities in Europe. The city has implemented a number of smart city initiatives that aim to improve the quality of life for residents and visitors, and to make the city more sustainable and efficient. It has Smart Energy, a smart grid that uses advanced metering infrastructure (AMI) to monitor and manage energy consumption. This helps to improve energy efficiency and reduce greenhouse gas emissions. Hamburg also has a comprehensive public transportation system that includes buses, trains, and ferries. The city

also has a bike-sharing program and is working on developing electric vehicle charging infrastructure. In addition, Hamburg has implemented some digital services and platforms that improve the transparency and efficiency of city government and has several initiatives to improve the health and well-being of residents, such as the development of a digital health platform that connects patients with healthcare providers and the promotion of active transportation through the construction of new bike lanes and the development of a bike-sharing program.

12. São Paulo, Brazil: São Paulo is the largest city in Brazil and the most developed in terms of smart city initiatives. The city has implemented a number of smart city projects, such as the installation of a network of environmental sensors to monitor air and water quality, and the creation of a mobile application that allows residents to report and track issues such as potholes and broken streetlights. The city also has a comprehensive transportation management system that allows residents to plan their routes, view traffic conditions and pay for parking.
13. Rio de Janeiro, Brazil: Rio de Janeiro has been working on a number of smart city initiatives, such as the implementation of a smart lighting system that adjusts the brightness of streetlights based on the amount of ambient light and a smart waste management system that uses sensors to monitor the fill level of trash bins and optimize garbage collection routes. The city also has a citizen engagement platform that allows residents to report issues and suggest improvements.
14. Recife, Brazil: Recife has implemented a number of smart city projects, such as the installation of a network of environmental sensors to monitor air and water quality, and the creation of a mobile application that allows residents to report and track issues such as potholes and broken streetlights. The city also has a comprehensive transportation management system that allows residents to plan their routes, view traffic conditions and pay for parking.
15. Florianópolis, Brazil: Florianópolis has been working on a number of smart city initiatives, such as the implementation of a smart lighting system that adjusts the brightness of streetlights based on the amount of ambient light and a smart waste management system that uses sensors to monitor the fill level of trash bins and optimize garbage collection routes. The city also has a citizen engagement platform that allows residents to report issues and suggest improvements.

These are just a few examples of smart cities in Latin America and Europe, but many other cities have implemented smart city initiatives, and this trend is likely to continue as technology and data become increasingly important for urban development in the regions.

4. Digital rights

Despite its undeniable prospect of offering a better life to citizens that habit Smart cities, most often, they are associated with practices that can harm digital rights and the exercise of human rights by collecting and using large amounts of personal data without proper consent or regulations in place (as pointed out by many documents published by the OECD[14]). This can lead to privacy violations and the potential for data misuse. Additionally, the increased use of surveillance technologies in smart cities can raise concerns about government overreach and the infringement on civil liberties. On the other hand, smart cities can benefit digital rights and human rights by providing citizens with more efficient and convenient services, such as improved public transportation and access to government services online[15]. They can also improve public safety and help address issues such as environmental degradation. Additionally, by leveraging data and technology, smart cities can help promote more equitable and inclusive outcomes for all residents[16].

Overall, smart cities need to be developed and implemented with a focus on protecting digital rights and human rights through the use of clear and transparent data policies and regulations, as well as robust oversight and accountability mechanisms[17].

Smart cities have the potential to greatly impact the future of education and upskilling in some ways, such as online and distance learning (by providing the infrastructure and technology necessary), personalized learning (by enabling the use of data and analytics to provide personalized learning experiences), workforce development (by upskilling and retraining the local workforce to meet the demands of new and emerging industries)[18]. Smart cities can also enhance the learning experience by utilizing virtual and augmented reality to provide immersive educational experiences and hands-on learning opportunities and investing in Smart libraries and community centers.

Furthermore, this scenario is encouraging Public-private partnerships, in a way to foster collaboration between government, education providers and industry to create education and upskilling programs that align with the needs of the local workforce and economy[19]. However, it is important to note that these benefits can only be achieved if the smart city is designed and implemented

14 See https://www.oecd.org/digital/privacy/.

15 Townsend, A. (2013) 'Smart Cities: Big Data, Civic Hackers, and the Quest for a New Utopia', *WW Norton & Co: United States.*

16 Ibid.

17 Crawford, S., Goldsmith, S. (2014) 'The Responsive City: Engaging Communities Through Data-Smart Governance', *Jossey Bass Publisher: United States.*

18 Ibid.

19 See https://www.mckinsey.com/capabilities/operations/our-insights/how-can-the-private-and-public-sectors-work-together-to-create-smart-cities.

in an equitable and inclusive manner and that access to technology and digital resources is provided to all citizens regardless of their socio-economic status[20]. In this manner, Smart cities are often seen as a potential solution to improve the delivery of public services, increase citizen engagement, and promote economic growth. On the other hand, it is wise to consider the potential negative ramifications of technology gaps, as urban centers require significant investments in infrastructure and digital provision, which can be a barrier for many cities and regions.

5. Final considerations: building a sustainable future with smart cities

The future of humanity is a multi-faceted and complex problem, and smart cities are just one piece of the puzzle. Ultimately, the success of smart cities will depend on their ability to balance the benefits and drawbacks of these technologies and to ensure that they are inclusive, equitable, and responsive to the needs of all residents[21]. The first step towards that is to ensure responsible and equitable data collection and usage from citizens to create and automate intelligent systems[22], preventing data extractivism practices.

We know that not only Smart cities rely heavily on the collection and analysis of data, as many current and future technologies have their predicament based on humanity's data[23]. It is, therefore, crucial that sensitive information, such as personal identification, location, and behavior, are treated in a responsible, equitable and legal manner. As such, it is noteworthy to comply with data protection regulations, such as the General Data Protection Regulation (GDPR) in the European Union, as well as other regional and national data regulations (for example, the Brazilian LGPD – Lei Geral de Proteção de Dados). To ensure compliance[24] with regulations, smart cities should implement data protection measures such as i) conducting data protection impact assessments to identify and mitigate any potential risks associated with data collection and processing; ii) providing clear and transparent information to citizens about the data that is being collected and how it will be used; iii) obtaining explicit consent from citizens for the collection and use of their data; iv) Implementing technical

20 Ibid.

21 Speck, J. (2013) 'The Walkable City: How Downtown Can Save America, One Step at a Time', *North Point Press: United States*.

22 Kramp, T., Kranenburg, R., Lange, S. (2013) 'Introduction to the Internet of Things' in: *et al. Enabling Things to Talk*, 'Springer, Berlin, Heidelberg'. Available at: https://doi.org/10.1007/978-3-642-40403-0_1.

23 Speck, J. (2013) 'The Walkable City: How Downtown Can Save America, One Step at a Time', *North Point Press: United States*.

24 See https://www.dataprotection.ie/en/organisations/know-your-obligations/data-protection-impact-assessments.

and organizational measures to protect data from unauthorized access, use, or disclosure; v) having a process in place for citizens to exercise their rights under data protection regulations, such as the right to access, correct, or delete their personal data.

When it comes to data from minors, smart cities should ensure that they have additional protections in place. They should have stricter controls on data collection and use and obtain explicit parental consent for the collection and use of data from children under the age of 18. In addition, they should also provide age-appropriate privacy notices and controls for children. It's worth noting that data protection regulations and best practices are continuously changing, so smart cities should regularly review and update their data protection policies and practices to ensure they are in compliance with the latest regulations.

In many developing countries, there is a significant digital divide, with a lack of access to digital technology and the Internet, particularly in low-income and rural communities. Smart cities have the potential to address this digital divide by providing access to digital technology and services for all citizens, regardless of their socio-economic status. They can work to solve the inequity involved in unequal access to digital technology in the Global South through:

1. Infrastructure development: investing in developing digital infrastructure, such as high-speed internet and wireless networks, it is crucial to ensure that all citizens have access to digital technology. This can include providing public Wi-Fi in public spaces, installing digital kiosks in low-income areas, and investing in community-based broadband projects.
2. Digital inclusion programs: developing programs to help bridge the digital divide by providing digital literacy training, computer access, and internet access to underserved communities. These programs can focus on educating citizens on using digital technology and the internet and providing resources such as computers, tablets, and mobile devices to those who cannot afford them.
3. Public-private partnerships: partnering with private companies and organizations to develop digital inclusion initiatives. These partnerships can provide funding, resources, and expertise to help expand digital access and opportunities for all citizens.
4. Digital governance: fostering digital technologies to improve governance, making it more inclusive, transparent, and responsive to citizens. This can include using digital platforms to provide citizens with easy access to government services and information and using data and analytics to improve service delivery and address the needs of marginalized communities.
5. Policy: by developing policies focusing on digital inclusion, such as providing subsidies for internet access to low-income households and establishing regulations to ensure that digital services are accessible to all citizens.

Smart cities are not a panacea for the digital divide, but if implemented to prioritize inclusivity and equity, they can be an important tool to help bridge the digital divide, particularly in the Global South. Unfortunately, although smart cities may be able to offer benefits such as improved quality of life, increased efficiency, and reduced environmental impact, they are not a solution to the digital divide. Yet, the digital divide is a complex issue that is rooted in a variety of socio-economic and cultural factors. Nevertheless, suppose smart cities are implemented to prioritize inclusivity and equity. In that case, they can be an important tool to help bridge the digital divide, particularly in the Global South. This can be achieved by investing in technology and infrastructure that is accessible to all, as well as by providing education and training opportunities to ensure that residents of all income levels and backgrounds have the skills they need to participate in the digital economy. Additionally, smart city initiatives should be implemented in a way that respects individual privacy and data rights. With a holistic approach, smart cities can make a meaningful impact in addressing the digital divide and creating more inclusive and equitable communities.

Chapter XIV
Communities' Governance in WEB3: the Role of DAOs

by María del Sagrario Navarro Lérida*

INDEX: 1. Web3: concept and relationship with blockchain. – 2. Value exchange model in Web3 ecosystems: uses cases. – 3. The role of DAOs on the Web3: the concept of DAO. – 3.1. Decentralized governance and value creation through tokenization.

1. Web3: concept and relationship with blockchain

The concept of Web3 has a philosophical connotation, some search for decentralization and democratization of the internet instead of control vesting in an oligarchic set of interdependent multinational corporations or traditional superpowers[1]. Web3 differs from Web 1.0 and Web 2.0, according to analysts, in both: the content and interactivity. Although the Internet has been in existence since the 1970s, the modern World Wide Web or Web 1.0 was tied directly to the launch of the graphical user interface beginning with the Mozilla (Netscape) web browser in 1994. In many ways, the Internet of the 1990s was a digital publisher and library, dominated by content owners digitizing, organizing, and pushing out content to the public in a one-to-many model. Web 2.0 refers to worldwide websites which highlight user generated content, usability, and interoperability for end users. Web 2.0 is also called the participative social web[2].

For advocates of Web3, Web 2.0 shifted from its participatory roots into a centralized, algorithmically mediated new media marketplace.

---

* Tenured Professor of Commercial Law UCLM. Member of Academic Advisory Body INATBA. Advisor in Internet Native Organization - Estonia-. Visiting Professor Universitá degli Studi (Milan). Author of more than 50 publications. Research stays in Bonn, Turin, Warsaw and Harvard. Since 2017, my field of study and work is mainly Blockchain, and in particular, Governance issues - DAOs, and DeFi-, Blockchain for good, and sustainability. A version of this paper is published in Digital Law and Innovation Review, ISSN-e 2659-871X, No. 16 (April-June), 2023.

1 Tan, J.Z., Langenkamp, M., Weichselbraun, A., Brody, A., Korpas, L. (2024) 'The constitutions of Web3'. Available at: https://arxiv.org/abs/2403.00081.

2 Garon, J.M. (2022) 'Legal Implications of a Ubiquitous Metaverse and a Web3 Future', *Marquette Law Review*, vol. 106(1). Available at: http://dx.doi.org/10.2139/ssrn.4002551.

– Web 2: Users are the data, corporations own the platform, and the code is closed.
– Web 3: Users own their data, contributors own the platform, and the code is open.

More recently, the development of blockchains and other cryptographic protocols have enabled the rise of a series of services and platforms collectively referred to as Web3. So, the term Web3 was coined in 2014 by Gavin Wood, Ethereum co-founder, who described it as a "decentralized online system based on blockchain". Several years later, crypto investor Packy McCormick called it "the internet owned by the builders and users, orchestrated with tokens". It is often difficult to draw the line between web2 using crypto and web3. What degree of decentralization does a web3 service need? How should the governance be organized[3]?

These technologies revolve around trustless mechanisms for social, economic, and political coordination. Many proponents of Web3 like to emphasize that typical forms of human governance (such as votes and constitutions) are not necessary or even desirable in Web3 communities. Other proponents of Web3 emphasize the importance of maintaining human-human mediation in addressing coordination problems. Despite the range of perspectives on how governance is best done, it is clear that group coordination requires shared goals and values[4]. Blockchains have enabled innovation in distributed economic institutions, such as money (e.g. cryptocurrencies) and markets (e.g. DEXs), but also innovations in distributed governance, such as DAOs, and new forms of collective choice. The "romantic" view of blockchain governance is collective choice and consensus through community voting. The exchange view, instead, is focused on entrepreneurial discovery of opportunities for value creation in governance space through innovation in protocols (e.g. Curve, Convex, Lido, Metagov, etc) that facilitate exchange of coordination and voting rights, that are newly made possible through tools that enable pseudonymous, composable and permissionless governance actions. The exchange lens on web3 governance also helps illuminate how this emergent polycentric process can generate robustness in decentralised systems[5].

But what is blockchain? It is clear that in this work we cannot analyze the concept of Blockchain technology itself, about which the doctrine has already been profusely written, but rather how some of the applications of this technology, hand in hand with smart contracts, allow us to return to the analysis

3 See https://www.marketscreener.com/news/latest/What-is-web3-and-should-it-be-taken-seriously--42844255/.

4 See https://constitutions.metagov.org/article.

5 Potts, J., Allen, D. W.E., Berg, C., Lane, A. M., MacDonald, T. (2022) 'The exchange theory of web3 governance', *Kyklos*, vol 76(4), pp. 659-675. Available at: https://ssrn.com/abstract=4209827.

of the concept of "company", of the contract that is at its base and of how the conflicts of interest that occur within it find a new way of apprehending themselves in the chain of blocks. blockchains are systems of rules generated (and then evolved) through consensus and agreement. blockchain romanticism is tied in with the underlying ethos of Web3 governance, including ideals of democracy, community ownership and decentralization[6].

Since the blockchain can be conceived as a decentralized distributed database of verified transactions that take place through the P2P network system and that operates on cryptographic algorithms, its value as an effective mechanism of disintermediation of the economy cannot be denied, with a potential undeniable. Thus, a new economic model is glimpsed, which can sometimes be understood as an advanced prototype of collaborative economy, and which is characterized by the tendency to decentralize hierarchical structures; a more responsive, transparent, and accountable approach to decision-making; and the inclusion of multiple interest holders in a dialogue platform, to find consensus-based solutions to common problems[7]. And there is the opportunity to mix the concepts of blockchain and Web3. In web3, governance is additionally an opportunity and site for entrepreneurship and discovery of value due to the capabilities and affordances of a range of new platform and protocol tools and institutions. The new governance tools in Web3 protocols brings into view the ways that collective choice infrastructure reveals and coordinates knowledge (including through coalition formation and vote delegation, buying and selling, which is to say just as in markets) while protecting against opportunistic behaviour. Blockchain and crypto are often considered to be the among the technologies that are most likely to usher in the Web3 revolution because they are designed to facilitate decentralized, permissionless, and trustless interactions.

Governance in Web3 is still a nascent field, as much an experimental art as a theoretical science. We can distinguish between two high-level views of governance:

1. the romantic vision of Web3 governance through consensus and collective action;
2. the exchange vision of governance as an emergent process of permissionless discovery and coordination over voting rights.

The exchange view is particularly useful for understanding innovation in Web3 governance. In this view, governance is part of a discovery process of rules for coordination and competition under uncertainty, in order to create value[8].

---

6 Ibid.

7 Atzori, M. (2015) 'Blockchain Technology and Decentralized Governance: Is the State Still Necessary?', *Journal of Governance & Regulation*, vol. 6. Available at: http://dx.doi.org/10.2139/ssrn.2709713.

8 Potts, J., Allen, D. W.E., Berg, C., Lane, A. M., MacDonald, T. (2022) 'The exchange theory of web3 governance', *Kyklos*, vol. 76(4), pp. 659-675. Available at: https://ssrn.com/abstract=4209827.

Well, DLT technologies – blockchain –, is an optimal instrument from the perspective of creating and sharing value. Indeed, blockchain technology also offers, and as the last point of analysis of this paper, alternatives to corporate governance problems, and in this sense, the so-called DAOs – decentralized autonomous organizations – change the nuclear opinion that governance needs to conceptually pivot on the agency theory of a community.

2. Value exchange model in Web3 ecosystems: uses cases

How can blockchain help in the task of delimit the way in which decisions are made in the network? How do you participate equitably in the value of the network? How and who should be responsible for what is done in the network? This is where, taking a further step, we can ask ourselves if perhaps a way to properly understand distribution network relationships, to correctly grasp integrated distribution in this new global and digitized economy, can go through the concept of "decentralized autonomous organization" (DAO), as one of the so-called pivots in the world of Blockchain. So, we can ask ourselves how a value chain should be and be managed in a decentralized reality in Web3.

In Web3 is important to talk about the possibility of the crypto economy, tokenization and decentralized governance to achieve "firms" in the deep sense of the term, communities with a common interest, more sustainable value chains. The value creation model in the crypto economy is therefore very interesting, to try to affect the real economy from it. Web3 talks about the creation of communities in which the generation of value and the distribution of that value, hand in hand with blockchain and tokenization, is possible. A more equitable mode of governance in which the different stakeholders participate.

In short, we are talking about the drift towards what is known as decentralized societies. We are talking about communities as networks. Networks with increasing returns are most efficient when treated neither as purely public nor purely private goods, but rather as partial and plural shared goods. DeSoc provides the social substrate to unbundle and recongure rights – rights of use ("usus"), rights to consume or destroy ("abusus"), and rights of prot ("fructus") – and enable ecient governance mechanisms across these rights that augment trust and cooperation while checking for collusion and capture[9].

In the field of the Web3, therefore, a crypto-economic model emerges. As we have already pointed out, unlike Web 2, where users are the data, corporations own the platform, and the code is closed, in Web 3 users own their data, contributors own the platform and the source is open.

9 Weyl, E.G., Ohlhaver, P., Buterin, V. (2022) 'Decentralized Society: Finding Web3's Soul'. Available at: https://ssrn.com/abstract=4105763.

Based on these conceptual premises, new value-generating ecosystems in which a plurality of actors interact are proliferating, hand in hand with blockchain and tokenization. And it is here where the figure of the DAO emerges as a catalyst, as we will see. These ecosystems adopt the "to earn" formula, prioritizing different concepts depending on the area in which they operate and its tokenomic design. The challenge for the lawyer is to understand the specifics of each model, but mainly to apprehend what they have in common: a decentralized governance system that creates incentives for its participants.

Some models:

I. Play or game to earn

When Axie Infinity launched in 2018, play-to-earn gaming became a sensation. The game turned into a viable source of additional funds for numerous low-income families in the Philippines. However, as Axie Infinity grew and took over the global blockchain-gaming space, the rewards ratio quickly started dwindling.The main reason for this significant cut in returns for players hides in the design of the game itself. One of the major challenges play-to-earn projects face is creating a self-sustainable in-game economy. Currently, most P2E games start out strong and offer lucrative rewards to players. However, increased in-game token issuing, as more players join, leads to significant value reductions. This, in turn, diminishes players' earning potential[10]. A service economy approach in the traditional sense of the concept means that users who have financial resources but lack time, pay other users to perform tasks in return for a reward. This concept can easily be translated into an in-game economy setting, where players are rewarded for completing tasks but also can delegate these tasks to others and share the rewards with them.

Incorporating the service economy approach brings a lot of utility to Web3 gaming projects. As money-rich players dedicate tasks to time-rich players, the game becomes a lot more sustainable without the need for developers to constantly offer input and cheap upgrades. This type of in-game economy is predicted to bring forward the next generation in blockchain-based gaming. Web3 is the internet with true ownership, as it provides a built-in layer that makes it easy to hold and transfer value. For sports leagues and their fans, Web3 can be a game changer in building direct relationships, aligning incentives and enabling true ownership and influence. Web3 gaming is all about giving players ownership of their in-game digital items and using tokens to create in-game economies (and potentially, game governance). Players with a stake in the game they're playing, whether through in-game

10 See https://cointelegraph.com/news/a-new-generation-of-p2e-games-is-here-and-its-open-to-everyone.

items or quasi-equity (tokens), will feel more invested in the game's success and spend more time and money playing. If you hit critical mass amongst players, they'll want to buy and sell items (NFTs), and an in-game economy can emerge that shifts the traditional gaming business model[11]. An example is Kickstarter, which a few months ago opened a funding round for Chips, a game that rewards you for positive behavior – in video games, kind behavior is not rewarded as often as violence. The interesting thing is that the game includes 2 currencies: Chip Coin and Crumb Token. Crumb Token is a monetary asset in the game. It can only be obtained through the game and players can be rewarded for completing mini games. Chip Coin on the other hand is the governance token in the Chips metaverse. Its operation is tied to a DAO.

II. Learn to earn

The learn to earn model brought about by blockchain differs depending on the blockchain project. But generally, it refers to the acquisition of knowledge along with incentivizes for various learning activities. These rewards may vary based on the platform but are mostly in the form of cryptocurrencies – tokens- . An example is CoinMarketCap's "learn crypto, earn crypto" program, where learners can earn various crypto assets, from BNB on the Binance Smart Chain to $LIKE, Only1's native token. Other such programs include the Coinbase Earn Program, which allows Coinbase users to get free cryptos by acquiring crypto knowledge, and Singapore-based Phemex's educational program called "Learn and Earn". Learn to earn programs are a great asset, especially for crypto beginners, as they provide educational resources while distributing tangible incentives to those involved. And, perhaps most importantly, they also pave the way for the democratization of education as learners don't need any particular qualifications or previous knowledge to get started with L2E. Models in fact that are aligned with the need for training in "crypto" as a strategy to also protect the participant in these ecosystems.

III. Vote to earn

As an example, the KAIF Holding DAO project was interesting, although it did not go ahead. The project was presented as the world's first Vote-To-Earn platform for the real and virtual world that allows users to earn money by voting or making proposals for big brands in any field (sports, entertainment, restaurants, food, etc.) . Thus, the football club DSV Leoben used the platform to interact with its fans. In addition to voting, the club encouraged

11 See https://resources.messari.io/pdf/messari-report-crypto-theses-for-2023.pdf.

users to complete activities (take selfies with players or use club merchandise) and submit proposals that create incentives for its participants.

IV. Healthcare

VITADAO, VIBE DAO, GENOMIC DAO: one DAO is leveraging Web3 technology to accelerate longevity research. Recognizing that age-related diseases strain healthcare resources, VitaDAO is centered on healthy life extension studies. Longevity is often promoted as a luxury that's accessible only to the wealthiest in society, triggering thoughts of Silicon Valley billionaires' intent on cheating death, but VitaDAO endeavors to stop income inequality from determining age-related diseases. The DAO is an accessible cooperative that anyone can join, working to guarantee that the future of longevity therapeutics is open to all. A human-centered approach is fundamental to the new wave of healthcare DAOs. For example, Vibe Bio distinguishes its organization from large profit-driven pharmaceutical companies by focusing on people over profits. The DAO empowers patients with rare and neglected diseases to go in search of their cure via community-governed research that discovers, funds, and tests promising treatments.

This approach expands the possibilities of treatment innovation, as Vibe Bio Co-founder Josh Forman tells: "Too often the best science and medicines never receive funding. Not because the science is bad, but because traditional systems optimize for different metrics – for example, prestige in academia, the lowest financial risk in biotech and influence in government. DAOs and decentralized science (DeSci) aim to empower researcher and patient communities to develop medicines and technologies that are potentially life-saving based on impact". Another interesting project is Genomic DAO, that seeks to break the centralization of R&D in modern medicine.

The design revolves around three interest groups:

i) Communities that initiate, drive and "govern" precision medicine research;

ii) research groups that will promote lines of research based on real medical needs;

iii) a direct-to-consumer marketplace for end products, including genetic insights, personalized recommendations, and precision medicine drugs.

In GenomicDAO, those who contribute investment or genetic data – the DAO token holders – are the ones who participate in the ecosystem value.

V. DeSci (Decentralized Science)

As said, Web3 distinctly differs from Web1 and Web2 from the prospect that a distributed network is established on blockchain, cryptocurrencies, and DAO to empower users with the ownership of the network. So first, from the perspective of economics, DeSci uses Web3 technology and open-source financial tools to introduce science and its services as the asset to the

market, such as the tokenization of intellectual property (IP), democratic governance of scientific systems, peer review, and data access. Second, from the perspective of organizational structure, DeSci is viewed as a set of mechanisms for bottom–up individual sense making. Individuals in DAOs can autonomously understand the world by defining problems, languages, and methods. Third, from the perspective of scientific management, DeSci aims at reforming the organization of scientific activity and improving the ability of science to fulfill its mission. DeSci revolutionizes the structure, norms, incentives, and value allocation of centralized scientific systems. DeSci is a new development paradigm built on decentralized technology protocols and organizational structures, such as Web3 and DAOs[12].

Currently, the main applications of DeSci are research for funding, knowledge sharing, and exploring scientific systems' ownership and value systems, such as decentralized funding, peer review, incentives, and domain-specific applications.

VI. Act to earn

For years I have been studying how Blockchain technology impacts the governance of companies. The importance of "governance" in sustainability – the G, of ESG criteria – is notable, although it is true that environmental and social issues have had more experience.

Sustainable governance has become one of the objectives of the EU[13]. Beyond issues that are also relevant - such as the concept of value chain in the field of business networks, the change of this concept towards that of "chain of activities" and the responsibility of the "head of the network" for the damages that, in the field of sustainability, occur in that chain and the transition from soft law to an enforcement more along the lines of tort law-, the ESG DAO project, although not operational today, allowed us to appreciate the opportunity that blockchain and tokenization have.

The project aimed to create an "open, democratic and neutral ESG scoring system that will drive a new wave of Web3 applications to engage consumers and incentivize companies to create positive change in the world". Not a minor objective considering the difficulty of defining what the ESG criteria should really crystallize into. Now, the most interesting thing about the project was the ecosystem that it intended to create using decentralized governance and tokenization. In addition to the base protocol, the project

12 Ding, W., et al. (2022) 'DeSci Based on Web3 and DAO: A Comprehensive Overview and Reference Model', *IEEE Transactions on Computational Social Systems*, vol. 9(5), pp. 1563-1573. Available at: 10.1109/TCSS.2022.3204745.

13 Directive (EU) 2024/1760 of the European Parliament and of the Council of 13 June 2024 on corporate sustainability due diligence and amending Directive (EU) 2019/1937 and Regulation (EU) 2023/2859.

was based on a cryptoeconomic design that encourages the participation of interested parties, creating for example an Expert Dao Council, while incorporating positive externalities, rewarding the involvement of the parties through "reward tokens" within the framework of "act-2-earn" designs.
Well, what do these business models that are appearing in the Web3 space have in common? The creation of a community involved in a project, which participates in the creation of ecosystem value and obtains benefits through a token economy model.
And to "organize" that community, the phenomenon of DAOs is used. Blockchain-powered tokens and shared ownership address the fundamental issue with centralized networks: value is accumulated by a single organization, which then conflicts with its own stakeholders. In addition, data independence is guaranteed by Web3 DApps using blockchain technology[14].

3. The role of DAOs on the web3: the concept of DAO

We cannot in this paper analyze the figure in depth[15], but it is enough to remember here that DAOs must be understood as a set of smart contracts that enable the organization and management of a "company" – understood as community – hosted on blockchain. So, a DAO is a decentralized organization deployed on a blockchain, which formalizes and automates its governance rules using software. Being implemented on a blockchain, the decentralized decision-making process and management function is coded directly into the software, through smart contracts. To create a DAO, a few elements are needed, among them: a purpose, a voting mechanism, a governance token, creating a community involved with the project and a mechanism to manage funds (treasury). The creation of a DAO requires the development of a governance structure to define roles and responsibilities for the participants in the network. The governance structure should address basic questions about entry and exit criteria, participants, transaction validation, role and responsibility of nodes, business logic change management, dispute resolution, etc. From the point of view that concerns us, that Blockchain network should allow to align the incentives with the goals of the participants keeping the incentives relevant to the needs of the participants over time and establish a defined process to through which the participants can decide on future changes in the governance model. And that governance model is what DAOs can offer.

14 We have analysed the relationship between DAOs and sustainability in https://internetnative.org/dao-phenomenon-improve-sustainability-standards/.

15 Navarro Lérida, M. S. (2018) 'Gobierno corporativo, blockchain y smart contracts. Digitalización de las empresas y nuevos modelos descentralizados (DAOs)', *Revista del Derecho del Mercado de Valores*, vol. 23.

The DAO phenomenon faces many challenges: its legal personality[16], its liability[17] or the nature of the governance tokens.

It is precisely this issue that deserves analysis.

3.1. Decentralized governance and value creation through tokenization

We have said that Web3 creates ecosystems based on decentralized governance that seeks to create tokenomic systems that allow interest groups to generate value and see their contribution to the ecosystem rewarded. Well, since the DAOs are an important element of this design, the question to analyze is what type of token the governance tokens are.

The first issue is to understand what a token is and what type of tokens can be used within the framework of the Web3 ecosystems that use DAOs.

A token, understood as crypto – asset, means a digital representation of value or rights which may be transferred and stored electronically, using distributed ledger technology or similar technology. There are different types of tokens – utility token, security token, payment token, non-fungible token (NFTs), stable coins, asset-backed tokens. And there are also different technical standards attached to these types. DAO governance tokens have traditionally been assimilated to utility tokens, understood as a type of crypto-asset which is intended to provide digital access to a good or service, available on DLT, and is only accepted by the issuer of that token. But lately the DAO tokenomic design includes more options. Por example, "traditional" token and NFT combination[18]. In short, the question has to do with the "rights" that are tied to the conditions of token holder in the DAO.

So, the second issue that deserves reflection is the idea of "rights" of token holders. DAOs are said to have the potential to eliminate the idea of a unified decision-making center in favor of true decentralized governance, eliminating the problem of agency theory that occurs in the framework of

16 Thus, regulation is proliferating that seeks "legal wrappers" for DAOs. On this matter, you can see Brummer, C. J., Seira, R. (2022) 'Legal Wrappers and DAOs'. Available at http://dx.doi.org/10.2139/ssrn.4123737; Mienert, B. (2021) 'How Can a Decentralized Autonomous Organization (DAO) Be Legally Structured?', *Legal Revolutionary Journal LRZ*. Available at: https://ssrn.com/abstract=3992329.

17 The aphorism that is at the base of the Blockchain philosophy, and therefore, of the DAO is Code is Law. Well, this aphorism does not seem to be satisfactory for the interests at stake. Cases like Olympus DAO, or more recently, the CFTC complaint against Ooki DAO and the token holders who exercise their vote in it, are the clear example that there is a lot of legal construction work to be done.

18 As it happens for example in the case of Optimism DAO with use two different types of tokens: one ERC20 (traditional token) and one ERC721 (NFT). This method is good for DAOs that want to create a small working group governed by different parameters from the larger DAO (think of a product team that's focused on shipping and needs full control over the vision) and DAOs that want to have decisions made in separate "houses" governed by different parameters.

corporate governance – although this elimination is not totally real – Regarding the first question, the debate is in the characterization of governance tokens, as said, which allow the members of a DAO to participate in its decision-making through a previously designed voting mechanism – not necessarily majority, but, for example, close to quadratic voting or futarchy – and in its nature or not as an "investment contract" in the terms that the SEC itself analyzed as a result of The DAO case or more recently that of American Cryptofed DAO, the first DAO constituted in Wyoming. Well, perhaps the nature of those tokens depends, as we have seen on how the DAO's "tokenomic" model has been designed. As an example, we could think of a design of incentives and rewards in Web3 models based on the greater involvement of the holder, thus creating non-expendable holder positions, close to "soulbounds tokens"[19], or we could create tokens that, in effect, are very similar to a security. Depending on the specific nature, in addition to the governance token, we would enter the scope of the MICA[20] regulation or not.

DAOs represent a progress compared to our Web2 systems. The DAOs are an essential element for the development of Web3, of the ecosystems that are being created based on decentralized governance and the generation of value. However, there are still many challenges and issues to analyze and clarify.

19 Concept created by Weyl, E. G., Ohlhaver, P., Buterin, V. (2022) 'Decentralized Society: Finding Web3's Soul'. Available at: https://ssrn.com/abstract=4105763. The concept of soulbound token refers to bound to an address, non transferable and thereby non tradeable. Are meant to represent something purer, and more tied to the identify of contributors. They could make governance and rewards more equitable an "fair" than mechanims with tradeable tokens. "The future of property innovation is unlikely to build on wholly transferable private property so far imagined Web3".

20 See https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52020PC0593&from=ES.

Chapter XV
Revolution of Contract Through Legal Technologies: Current Trends in Contract Automation

by Silvia Martinelli[*] and Carlo Rossi Chauvenet[**]

INDEX: 1. Introduction. – 2. Current trends in contract automation. – 3. Conclusion.

1. Introduction

The contract in its unlimited and inexhaustible possibilities of adaptation seems eternal, just as the need for cooperation between men is eternal, however, contract law is not eternal but transitory, and so are the values it contains[1].

The broad use of software, data, algorithms and AI has shaped many economical activities and introduced new possibilities/tools for the management of resources. Big data enable a more granular vision of the world and the ability to collect, process, analyse and get insights from this information,

---

* Research Grant Holder at the University of Turin and she obtained the Italian National Qualification for the role of Associate Professor in Private Law.She wrote two monographies: "I contratti della platform economy. Ruoli e responsabilità delle piattaforme, Giappichelli, 2023" and "Diritto all'oblio e motori di ricerca. Memoria e privacy nell'era digitale", Giuffrè, 2017 and many scientific articles about e-commerce, data protection, private law, provider's liability and platform economy, contracts. Winner and PI of the JM Module "PLATFORMLAW - Platform & Data Economy European Legal Framework: platform, data-driven business models, AI and contract automation", ERASMUS-JMO-2021-HEI-TCH-RSCH, www.platformlaw.unito.it. Lawyer and Strategic Research Manager in Data Valley Consulting.

** Managing partner of a law firm specialized in IT Law, Company and Privacy Law is a Data Protection Officer (DPO) and an expert in legal tech issues related to digital business models. Carlo is Adjunct professor at Bocconi University and at the University of Padova. He is coordinator of the Legal Clinic of the startup accelerator "Bocconi4Innovation", co-founder of "Sweet Legal Tech", consultancy company in Legal Tech, and co-founder of "Iubenda", the first Italian Legal tech company. Author of numerous articles and publications he participates as speaker to international conferences also in his capacity of Chair of the "National Centre for IOT and Privacy" and manager of the "Data Valley" initiative, a program dedicated to the development of partnerships between PMI and digital Multinational companies.

1 Alpa, G. (2012) *Le Stagioni Del Contratto*. Il Mulino, which in turn quotes and re-elaborates Francesco Santoro-Passarelli, in the conclusion of his monography dedicated to the analysis of the evolution of contract and contract law during the years.

with incremental levels of automation. The digitalization of companies and the emergence of new business models data-driven and platform based, has reshaped business and contract management.

The broader phenomenon of servitization and the transformation of products and services into products "as a service"[2] therefore also has an impact on contracts and relationship management.

The expression "legal tech"[3] identifies software and innovative solutions – both with reference to all existing legal professionals (and not only), and with respect to the individual activities that the jurist performs – which aim to either solve specific problems or to eliminate inefficiencies by organizing activities in a new way, adopting solutions that new technologies enable[4]. Legal work "can be improved, becoming cheaper and more accessible for a greater proportion of population"[5]. Lawyers' role in the ecosystem can be on the side of users, as customers or consultant, or on the side of the legal tech company, working in the creation and evolution of the software.

Contracts are part of this change[6]. No longer represented by paper documents or pdf, but with a digital model made of orders – state of work (stable) and terms and conditions or framework agreements (modifiable during the relationship).

---

2 Susskind describes the evolution of legal services as a 'mercification': the market moves from a personalised and tailored consultancy to a standardised and automated service available online. the evolution of legal services as a 'mercification': the market moves from a personalised and tailored consultancy to a standardised and automated service available online. The expression 'as a service' is used to indicate the use of cloud applications to realise different functions and offer it to the final user as a software product. See Susskind, R. (2013) *Tomorrow's Lawyers. An Introduction to Your Future*. Oxford University Press. With regard to the legal service's evolution and 'marketisation and commodification of legal practice', see also Caserta, S. (2020) 'Digitalization of the Legal Field and the Future of Large Law Firms', *Laws*, vol. 9(2). Available at: https://doi.org/10.3390/laws9020014.

3 See also Hartung, M. (2018) *The Digital Transformation*, in Hartung, M., Bues, M.M., Halbleib, G. (eds)(2018) *Legal Tech. A Practitioner's Guide*. Beck; Messori, G. (2020) *Legal Tech*, in Ziccardi, G., Perri, P. (eds)(2020) *Dizionario Legal Tech*. Giuffré Francis Lefebvre.

4 See ibid, which identifies three steps to increase efficiency of existing legal products: 1) Identify the potential for increasing the efficiency of your key products and services; 2) Analyse the workflow process of the most important legal products (workflow analysis); 3) Implement necessary technical and organisational measures. See Caserta, S. (2020) 'Digitalization of the Legal Field and the Future of Large Law Firms', *Laws*, vol. 9(2). Available at: https://doi.org/10.3390/laws9020014.

5 Hartung, M., Bues, M.M., Halbleib, G. (2018) (eds) *Legal Tech. A Practitioner's Guide*. Beck.

6 See Schuhmann, R. (2020) *Quo Vadis Contract Management?*; Cummins, J., Clack, C. (2022) 'Transforming commercial contracts through computable contracting', *Computers and Society,* vol.1; Ebers, M., Poncibò, C., Zou, M. (2023) 'Contracting and Contract Law in the Age of Artificial Intelligence', *Bloomsbury Publishing PLC*; Martinelli, S., Rossi Chauvenet, C. (2022) *Legal tech, contract re-design & big data per professionisti e imprese*, Wolters Kluwer; Martinelli, S. (2023) 'AI as a Tool to Manage Contracts', *ERPL*.

The representation, the form of the contract, is not only the document but software design and data and the whole contract process or entire groups of contract can be managed and analysed together with a granular vision, to extract relevant information.

The contract is therefore moving from document to data, represented as a set of fluid information and it can be organised and managed by technology for the efficient (and peaceful) relationship of the parties, but also analysed to identify corporates' risks and compliance or develop new businesses.

Data enables new forms of information acquisition and management which can also be applied to improve contracts and the relationship between the parties[7], as continuous reference of business activities. Furthermore, the streamline offered by legal tech software can expand the number and categories of individuals able to directly draft and manage contracts.

The legal tech software for contract drafting, contract analytics, and contract lifecycle automation is changing the way we write, conclude, and manage contracts and they will do it even more in the next few years. The adoption of these technologies in the legal field brings with it some significant changes in the ways in which the legal activity is performed and impacts on contract practices and contract law.

The question that arises is how these changes will impact the contract and contract laws. This change is still understudied and the impacts it will have on relationships and contracts are yet to be understood. The contribution will be limited to highlighting the main ongoing trends.

2. Current trends in contracts automation

The observation of the ongoing phenomena shows a change in the subjects involved in the negotiation. The software fits deeply into the negotiations between the parties, facilitating contractual interaction thanks to the more efficient matching and profiling possibilities, but also standardization and customization. On the one hand, it therefore makes the conclusion and drafting of a contract for the user more accessible and easier, thus also widening the audience of recipients. On the other hand, it emerges as a new intermediary who, through the technological tool, facilitates and mediates the regulation of the relationship.

The new methods of collecting and using information certainly intervene in the formation of contractual will, even with the emergence of new risks. The simplicity of use and the user-friendly interface, possibly combined with AI

7 See Geis, G.S. (2008) 'Automating Contract Law', *N.Y.U. L. Rev.*, vol. 83. See also Van Erp, S. (2020) *Management as Ownership of Data* in Lohsse, S., Schulze, R., Staudenmayer, D. (2020) (eds) *Data as Counter-Performance – Contract Law 2.0?,* Hart Publishing. He underlines that when we use the word 'disruptive', "we really mean the impact of data, data processing, data analysis, data profiling and data transactions, put it short: the data economy".

assistants, intervene as facilitators but at the same time can lead to new deceptive practices and market inefficiencies. Furthermore, the acquisition of large amounts of reserved data by contract automation companies may pose additional concerns on the control of the relationship by these new intermediaries.

New rules may be introduced, regulating the transparency and independence of these new operators, as it already happens for financial and insurance intermediaries.

To understand the phenomenon in depth, a broader study of existing operators and available services would be needed. To this end, it is necessary first of all to distinguish between services offered to a large public (professional and consumers) and services provided tailored to individual companies with the integration of contract management tools into existing processes.

Moving from the subjects to the new possibilities enabled, a further relevant aspect in the ongoing change concerns the new possibilities offered by the data and the granular vision that they enable[8]. The software's capabilities to extract and analyse information can be used not only by companies but also by researchers, legislators and regulatory authorities.

"RegTech" is defined as "the use of technology, particularly information technology, in the context of regulatory monitoring, reporting and compliance" and it focuses on "the digitization of manual reporting and compliance processes"[9]. It could also "enable a close to real-time and proportionate regulatory regime that identifies and addresses risk while also facilitating more efficient regulatory compliance".

Firstly, focusing on companies, they can use the data and the software itself to align company's activities to the business goals they want to achieve, but also for compliance, to take the necessary steps to comply with laws, regulations, and policies.

Thanks to the usage of AI, we are moving from a reactive on demand model to a predictive full automated system of deployment of legal services. Compliance is the automatic result of the correct setting of a series flows, processes and triggers that automatically provides actions such as sending emails, letters, documents. The legal advisor becomes the "automator".

Secondly, the legislator or regulations authorities can look at these software and data to monitor the business or to introduce new rules, as an obligation of the legal tech company (i.e. to avoid the use of a certain clause or to introduce a new model).

---

8 Busch, C., De Franceschi, A. (2018) *Granular Legal Norms: Big Data and the Personalization of Private Law*, in Mak, V., Tjong Tjin Tai, E., Berlee, A. (eds) (2018) *Research Handbook on Data Science and Law*, Edward Elgar.

9 Arner, D.W., Barberis J.N., Buckley, R.P. (2017) 'The Emergence of Regtech 2.0: From Know Your Customer to Know Your Data', *SSRN Electronic Journal*.

Thirdly, the contractual data analysis may help researchers to better understand the contract, enabling new analysis of clauses and sectors' practices, the discover of best practices or legal inefficiencies and a more detailed comprehension of parties and relationships.

In addition to the new possibilities offered to jurists by data analysis, those offered by the new infrastructure in which information is offered also appear particularly interesting. The new organization of information structured in the design of the tool or platform may be also used by jurist for regulation.

Looking at the form of communication of the information, the software can help the weaker party in the comprehension: i.e. underling relevant information, requiring attention to a specific clause or comparing clauses, or giving simple or wide explanations accompanying the clause.

Furthermore, the by design regulation can be introduced, for instance regarding marketplaces information duties: i.e. if the party is a consumer the consumer's law can be assured with the automatic insertion of the standard information of consumer's rights or giving evidence if a clause seems vexatious.

A further interesting tool offered on contract drafting software is the possibility to give an evaluation of a clause, which can be represented through the use of colors or traffic light (green or red) or with a score, as in the reputational feedback system used in platforms. The evaluation can be internal, referred to the single company, or collective, with the suggestions of all the users of the software.

This could lead to an evaluation of the same legal production that becomes a product by its users. Qualified evaluators, such as consumer or trade associations or authorities, could also be included.

Moreover, using code it is also possible to introduce new rules on the software and it will apply the rule created to all the users, whether the particular condition described is verified. This is the normal functioning of coding but it can be used also to introduce legal rules. The legal rule, even that deriving from the contract and from individual autonomy, became part of the technology itself.

This form of regulation is considered in the European Regulation 679/2016, 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)". Article 25 of the Regulation is dedicated to the "data protection by design and by default" and it introduces an obligation for the data controller to implement "appropriate technical and organisational measures" and to "integrate the necessary safeguards into the processing" in order to meet the requirement of the GDPR and protect the right of the data subjects.

The "by design" principle, introduced by the European Regulation 679/2016 for data protection (GDPR) can spread in all areas of law, together with the

greater diffusion and capillarity of the new tools and in relation to the possibilities they offer.

3. Conclusions

The use of legal tech software to contracting, automate contract drafting, and contract management is changing contracts and it will impact contract law.

The major trends identifiable in this transformation concern i) the extension of the categories of recipients and the emergence of new intermediaries, ii) data analysis to acquire new information on contracts and companies and manage them more efficiently; iii) the regulatory possibilities offered by the software itself in relation to the design of the product itself and the possibilities of improving the communication of contractual information or even integrating rules and principles by design.

As it happens in other markets, platforms and software creators are emerging as new intermediaries in the relationship between users/parties, simplifying and governing the relations. The use of these software multiplies data processing and accessibility, also regarding contract and contract practices and these data, as well as the software/platform architecture itself, can be used to manage the relationship: monitor, analyse, regulate and govern contracts.

Chapter XVI
"Justice" on Digital Platforms: Internal Complaint-Handling Systems and Mediation in P2B Relationships. A Call for Reform

by Ludovica Sposini*

INDEX: 1. "Justice" for business users on digital platforms: an introduction. – 2. The internal complaint-handling system as the first way to solve conflicts in P2B relationships. – 3. When the in-house tool does not work: the mediation. – 4. Need for a change in perspective?

1. "Justice" for business users on digital platforms: an introduction

It is becoming increasingly evident that online platforms have assumed a key role in our everyday lives[1]. This is particularly true for professional subjects, who use digital platforms to run their business online in an attempt to expand

---

* PhD Candidate in Law at Sant'Anna School of Advanced Studies in Pisa, specializing in the regulation of digital platforms, consumer protection and P2B relations. She spent three months as visiting researcher at the European Legal Studies Institute (ELSI) at the University of Osnabrück under the supervision of Prof. Dr. C. Busch as well as at the University Carlos III of Madrid working with Prof. Dr. Teresa Rodríguez de las Heras Ballell. She was also nominated by the University of Pisa "Cultore della materia" in Private Law. Currently, she is a member of the Centre of Excellence "EURA" led by Prof. A. Bertolini working on the regulation of AI, Robotics and Advanced Technologies as well as a Research Fellow of the "ISLC" of the University of Milan. Furthermore, Ludovica is a Member Consultive Committee of the European Law Institute (ELI) working on several report and projects. Also, she is a Teaching Assistant at the University of Pisa, and she is doing an internship as Public Affairs and Legal Counsel at "Traent", a company developing hybrid blockchain technologies.

1 On the power of digital platforms, see Bughin, J., van Zeebroeck, N. (2017) *New evidence for the power of digital platforms*, The McKinsey Quarterly; Gawer, A. (2022) 'Digital platforms and ecosystems: remarks on the dominant organizational forms of the digital age', *Innovation*, vol. 24, pp. 110-124; Marty, F.M., Warin, T. (2020) 'Digital Platforms'Information Concentration: From Keystone Players to Gatekeepers'; Bamberger, K.A., Lobel, O. (2017) 'Platform market power', *Berkeley Tech LJ*, vol. 32, p. 1051; Hein, A. et al. (2020) 'Digital platform ecosystems', *Electronic Markets*, vol.30, pp. 87-98; De Reuver, M., Sørensen, C., Basole, R.C. (2018) 'The digital platform: a research agenda', *Journal of information technology*, vol. 33, pp. 124-135.

their activity and reach a greater number of consumers they would not otherwise interact with[2]. However, to achieve such result, business users are forced to turn to providers of online intermediation services and accept the contractual conditions they unilaterally impose on them[3].

Aware of the extent of the digital phenomenon and concerned about the weak and dependent position of business users towards platforms, the European legislator recently decided to face the problem directly, by adopting the Regulation (EU) 2019/1150 "on promoting fairness and transparency for business users of online intermediation services"[4] (henceforth P2B Regulation). This contribution will focus on the alternative resolution systems set up by the platforms themselves to resolve disputes arising during the contractual agreement. These tools are, indeed, more suited to the fluid and rapid relationships developed in the digital ecosystem, for at least two main reasons: on the one hand, they enable to solve the dispute in a very short time, and, on the other hand, they do not entail high costs for users. Thus, the importance of them in ensuring a full right of defense for business users is quite evident and, therefore, this reflection will examine alternative dispute resolution systems set up by platforms to verify whether they are effectively functional in providing a fair and equal digital environment or instead whether there is a need for a reconsideration.

---

2 See Karakas, F. (2009) 'Welcome to World 2.0: the new digital ecosystem', *Journal of Business Strategy;* Wirtz, B.W. (2019) *Digital business models*, Springer; Al-Debi, M.M., El-Haddadeh, R., Avison, D. (2008) 'Defining the business model in the new world of digital business', *AMCIS 2008 proceedings,* p. 300; Kotarba, M. (2018) 'Digital transformation of business models', *Foundations of management,* vol. 10, pp. 123-142.

3 Cauffman, C. (2019) *New EU rules on business-to-consumer and platform-to-business relationships*, SAGE Publications Sage UK: London, England, pp. 469-479; Graef, I. (2019) 'Differentiated treatment in platform-to-business relations: EU competition law and economic dependence', *Yearbook of European Law*, vol.38, pp. 448-499; Bostoen, F., (2018) 'Neutrality, fairness or freedom? Principles for platform regulation. Principles for Platform Regulation', *Internet Policy Review*, pp. 1-19. For an analysis of the weakness of business users in P2B relationships, see Cutolo, D., Kenney, M., (2021) 'Platform-dependent entrepreneurs: Power asymmetries, risks, and strategies in the platform economy', *Academy of Management Perspectives*, vol. 35, pp. 584-605; Kenney, M., Zysman, J. (2016) 'The rise of the platform economy', *Science and technology*, vol. 32, p. 61.

4 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 (2019). For a first comment, see Savin, A., (2022). Regulation 2019/1150/EU on Promoting Fairness and Transparency for Business Users of Online Intermediation Services, *EU Regulation of E-Commerce*. Edward Elgar Publishing, pp. 364-385; Smorto, G. (2020) *La tutela del contraente debole nella platform economy dopo il Regolamento UE 2019/1150 e la Direttiva UE 2019/2161* (cd Omnibus), Fairness e innovazione nell mercato digitale; Jablonowska, A. (2019) 'Regulation of online platforms in the digital single market', *Studia Prawnoustrojowe,* pp. 63-79.

2. The internal complaint-handling system as a first way to solve conflicts in P2B relationships

Internet service providers (ISPs)[5] are obliged under P2B Regulation first to set up an internal complaint-handling system to prevent disputes and, in any case, to resolve them in an efficient and timely manner[6]. With a view to effectively guaranteeing users' right of defense, any in-house system must be free of charge and easily accessible to all interested parties[7]. Not only that, but this system must also be based on two fundamental principles: the principle of equal treatment of business users and the principle of transparency[8]. It follows from the application of the first principle that business users are entitled to be treated equally in a complaint if there is an "equivalent situation". Mainly problematic, however, is the corollary that follows from this principle. In particular, if the ISP has to treat differently business users who are not on an equivalent situation, then it is also entitled to deal with the various issues "in a manner which is proportionate to their importance and complexity"[9]. On the other hand, the complaint is handled entirely by the ISP and, therefore, it will be the latter that decides when a dispute can be considered more or less "important" or more or less "complex". However, it must also be pointed out that the European legislator has not provided any further specification as to the criteria based on which the ISP makes such a decision, without even imposing an obligation to inform the user about them. In fact, in the absence of a legislative specification of these two concepts, the provider inevitably has a wide discretion in making its assessments with the risk, thus, of infringing the principle of equal treatment of business users.

---

5 We must point out that in this contribution we have used the terms 'ISP' and 'platform' in an atechnical way. For the differences between digital actors, see Bertolini, A., Episcopo, F., Cherciu, N. (2021) *Liability of Online Platforms*, Scientific Foresight Unit within the Directorate-General for Parliamentary Research Services (EPRS) of the Secretariat of the European Parliament, pp. 1-150.

6 Article 11 of Regulation (EU) 2019/1150. However, we must point out that the provisions provided by article 11 are not applicable to providers of intermediation services that are small enterprises within the meaning of the Annex to Recommendation 2003/361/EC.

7 See Can Pehlivanoglu, M. (2020) 'Internal Complaint-Handling Systems of Online Platforms, Commercial Law and the Board of Directors' Duty of Care', *J Marshall LJ*, vol. 14, p. 168; Conti, G. (2020) *Lineamenti di diritto delle piattaforme digitali. Le tutele del consumatore e dell'utente commerciale nei confronti dei cybermediary*, Maggioli Editore 2020, vol I, pp. 226-235.

8 Article 11 and Recital No. 37 of Regulation (EU) 2019/1150. For a comment, see also Busch, C. (2020), 'The P2B Regulation (EU) 2019/1150: Towards a "procedural turn" in EU platform regulation?', J*ournal of European Consumer and Market Law*, vol. 9, p. 1334; Regulation 2019/1150/EU on Promoting Fairness and Transparency for Business Users of Online Intermediation Services, pp. 364-385.

9 Article 11(1) of Regulation (EU) 2019/1150.

The further fundamental principle is that of transparency. According to this, the complainant is foremost entitled to receive a package of information on access, operation, and effectiveness of the in-house system and, specifically, the total number and main types of complaints lodged, as well as the time required to process them[10]. This information package must be prepared by the ISP already in its contractual terms and conditions, which the platform is subsequently obliged to make available to the public and update (at least) annually. However, the latter obligation exists only where "significant changes are needed"[11]. To ensure transparency in the substantive sense[12], the provider must not only inform individually the user about the outcome of the complaints procedure but must also ensure that he understands the mechanisms by which this system works. Despite the importance of this statement, the EU legislator merely imposed an obligation to provide the business user with a general description of the functioning of the system[13]. Its problematic nature is thus evident, because the fact that information is limited exclusively to a general overview of the functioning of the system runs the risk of resulting, in practice, in a mere list characterized by vagueness and excessively technical terms which the business user, in any case, is not capable of processing[14]. The ISP's decision following the initiation of the complaint procedure does not have judicial or even para-judicial effect but takes on the guise and effects of

10 Article 11(2)(3) of Regulation (EU) 2019/1150.

11 Article 11 (4) of Regulation (EU) 2019/1150.

12 On the principle of substantial transparency, see Pagliantini, S. (2019) 'I mutui indicizzati ed il mito di un consumatore "costituzionalizzato": la "dottrina" della Corte di Giustizia da Árpád Kásler a Dziubak', *Le nuove leggi civili commentate,* pp. 1258-1283; Minervini, E. (1997) 'La trasparenza delle condizioni contrattuali (contratti bancari e contratti con i consumatori)', *Banca borsa e titoli di credito*, vol. 50, pp. 94-111; Roppo, V. (2011) *Il contratto*, Giuffrè, p. 976; Di Bona, L. (2014) 'Spunti di riflessione in tema di obblighi informativi (e neoformalismo) nei contratti asimmetrici', *Studi Urbinati, A-Scienze giuridiche, politiche ed economiche*, vol. 65, pp. 229-266.

13 In this sense, see Conti, G. (2020) *Lineamenti di diritto delle piattaforme digitali. Le tutele del consumatore e dell'utente commerciale nei confronti dei cybermediary*, Maggioli Editore, p. 227; Lukovic, V. (2021) 'Information asymmetries in algorithms at digital platforms: motivations to participate and EU regulatory approach', E*MAN 2021–Economics & Management: How to Cope with Disrupted Times*, p. 167.

14 The behavioural sciences have amply demonstrated that an excessive amount of information can have an overload effect, as human beings are affected by some 'biases' that do not allow them to process all information. See Ben-Shahar, O., Schneider, C.E. (2011) 'The Failure of Mandated Disclosure', *University of Pennsylvania Law Review*, vol. 159, p. 688. For an in-depth analysis, Ben-Shahar, O., Schneider, C.E. (2014) *More Than You Wanted to Know: The Failure of Mandated Disclosure*, Princeton University Press; Simon, H.A. (1990) *Bounded rationality. Utility and probability*, Springer, pp. 15-18; Kahneman, D. (2003) 'A perspective on judgment and choice: mapping bounded rationality', *American psychologist*, vol. 58, p. 697; Gigerenzer, G., Goldstein, D.G. (1996) 'Reasoning the fast and frugal way: models of bounded rationality', *Psychological review*, vol. 103(4), pp. 650-669; Thaler, R.H., Sunstein, R.C. (2012) *Nudge: The Final Edition*, Penguin, p. 371.

a private agreement, insofar as it finds its justification in the light of the contract concluded between the provider and the business user[15]. In other words, the platform is legitimized in its role as decision-maker of the complaints procedure when the business user accepts the terms and conditions unilaterally prepared by the provider. From this it follows that, although the European legislator itself affirms the need to achieve a "quick and effective bilateral resolution", it is clear that the decision is taken unilaterally by the ISP (by virtue precisely of that "mandate" given to it by the user when entering into the contract) and that, however, it produces effects between the parties as if its content had been established by mutual agreement[16]. Neither is it possible to ensure full impartiality, especially where the business user has activated the internal procedure to complain of an infringement committed by the same party – the provider – who will then have to decide the matter[17]. The business users are certainly in a position of dependence that is not only economic, but we could say "existential", given that without the intermediation services offered by the platform there is a strong probability that their business would not even survive. In this context, there is not only the danger that the user finds himself disincentivized to activate a complaint process for fear of possible retaliation by the provider (who might decide to penalize the user's products or services through, for example, a lower position ranking), but there is more. The platform is entitled, by express legal provision, to keep the opposed measure in force, with no possibility for the business user to request its suspension even during the pendency of the complaint procedure. Such a provision may in fact represent a further disincentive to attempt the complaint procedure, because the business user bears a considerable injury for the entire duration of the procedure that often is a "reputational damage", which is difficult to quantify and for which an amount of money does not appear to be sufficient. It is for this reason that the European legislator has provided that, in any case, the attempt to find an agreement through the in-house system is without prejudice to the right of both parties to initiate a judicial or extrajudicial action at any time. In other words, since the platform is entitled to keep the measure harming the user's interests in force, the complainant could bring an action before the judicial authority to obtain an interim measure[18].

15 Conti, G. (2020) *Lineamenti di diritto delle piattaforme digitali. Le tutele del consumatore e dell'utente commerciale nei confronti dei cybermediary*, Maggioli Editore, p. 226.

16 Recital No. 37 of Regulation (EU) 2019/1150.

17 See Busch, C. (2020) 'The P2B Regulation (EU) 2019/1150: Towards a "procedural turn" in EU platform regulation?', *Journal of European Comsumer and Market Law*, vol. 9, pp. 133-134; Savin, A., Lodder, A., Murray, A. (2022) 'EU Regulation of E-Commerce EU Regulation of E-Commerce: A Commentary, Chapter 11: Regulation 2019/1150/EU on Promoting Fairness and Transparency for Business Users of Online Intermediation Services', *Edward Elgar Publishing*, pp. 364-385.

18 Recital No. 37 of Regulation (EU) 2019/1150. See Katsh, E., Wing, L. (2006) 'Ten years of online dispute resolution (ODR): Looking at the past and constructing the future', *U*

However, this provision may not actually be adequate to guarantee the protection of the business user because, on the one hand, the slowness of judicial proceedings is well known and, therefore, there is a serious danger that the judge may order the interim measure when the complaint procedure is already concluded (since there is no provision for suspending that process when the judicial authority is seized); on the other hand, bring the case to the court could be very onerous for the business user, not only in economic terms but also because some national legislations require particularly stringent criteria in order to obtain such a measure.

3. When the in-house tool does not work: the mediation

Where the internal complaint-handling system fails, it is possible to make use of a further alternative dispute resolution tool, which, for P2B relationships, is called "platform-to-business mediation"[19]. This is a new form of mediation insofar as the traditional one is only applicable to consumer relations and, therefore, essentially excludes the possibility of extending its rules to issues arising during the implementation of P2B relations[20]. Looking at platform-to-business mediation, it should first be noted its voluntary nature, and, in fact, recital No. 12 of the P2B Regulation simply states that "providers of online intermediation services should facilitate mediation". Precisely to make it as easy as possible for the parties to settle the matter without the need to go to court, a particularly deformalized procedure is also constructed here.

An obligation incumbent exclusively on the provider, however, is to specify in its terms and conditions at least two mediators[21]. The mediator, according

Tol L Rev, vol. 38, p. 19; Mania, K. (2015) 'Online dispute resolution: The future of justice', *International Comparative Jurisprudence*, vol. 1, pp. 76-86; Busch, C. (2020) 'The P2B Regulation (EU) 2019/1150: Towards a "procedural turn" in EU platform regulation?', *Journal of European Comsumer and Market Law*, vol. 9.

19 Article 3(a) of Directive 2008/52/EC of the European Parliament and of the Council of 21 May 2008, on certain aspects of mediation in civil and commercial matters defines mediation as: 'a structured process, however named or referred to, whereby two or more parties to a dispute attempt by themselves, on a voluntary basis, to reach an agreement on the settlement of their dispute with the assistance of a mediator'. In Regulation (EU) 2019/1150, see Recital No. 40.

20 The regulation of consumer ODR systems can be found in European Parliament and Council, Regulation (EU) No. 524/2013 of the European Parliament and of the Council of 21 May 2013, on online dispute resolution for consumer disputes and amending Regulation (EC) No. 2006/2004 and Directive 2009/22/EC (Regulation on consumer ODR)(2013). Article 2(1) of Regulation (EU) 2013/524 said that: "This Regulation should apply to the out-of-court resolution of disputes concerning contractual obligations stemming from online sales or service contracts between a consumer resident in the Union and a trader established in the Union".

21 As with the in-house complaint-handling system, online service providers that are small enterprises within the meaning of the Annex to Recommendation 2003/361/EC are exempt from the obligation to indicate mediators.

to Article 12(2) of the P2B Regulation, is: i) impartial and independent; ii) he provides his service at affordable prices for business users and without undue delay; iii) he is physically or even telematically reachable; iv) he knows both the language of the contract in force between the business user and the provider and the business-to-business relations themselves.

Particularly important and problematic are the qualities of impartiality and independence[22], because the choice as to the identity of the mediators is left to the platform (the entity in a position of strong dominance over business users) which could appoint individuals who are easily controlled or influenced. Nor is this risk mitigated by the desire in recital No. 40 that business users and the provider should nevertheless remain free to jointly appoint a mediator of their choice. Firstly, this provision is contained within a recital and thus certainly has a different prescriptive force from that of Article 12, and it must therefore be doubted whether the former is sufficient to guarantee the independence of mediators[23]. In such a situation, then, it is doubtful that the business user, definitely the weaker party, can actually make a free choice without any external pressure. Not only that, but even where there is a conflict, it cannot be taken for granted that the seller has the power to actively discuss with the platform and oppose to its mediator proposal, given that he may not even have the same information as the platform. However, one of the most problematic aspects of mediation concerns the effects and, above all, the enforcement of the agreement reached at the end of this procedure, as it does not have the same force as a court order. In this context, then, what happens if the agreement is not respected by one of the two parties and, especially, by the platform, provided the "non-binding" nature of it?

The P2B Regulation does not offer any type of sanction and Directive 2008/52/EC (which is often referred to), instead, delegates this task to the MS[24]. This has a series of negative effects on the principle of the effectiveness of the protection of the rights of commercial users, since the only available way is to take a legal action, with an inevitably increase in both time and costs as well as a

22 It is also permitted to appoint mediators who provide their services outside the EU, provided, however, that the business user is ensured all the guarantees deriving from European law and the laws of the MS. See article 12(1) of Regulation (EU) 2019/1150.

23 See Zlatanska, C., Betancourt, J.C. (2013) 'Online Dispute Resolution (ODR): What is it, and is it the Way Forward?', *Arbitration: The International Journal of Arbitration, Mediation and Dispute Managemen*t, vol. 79.

24 See recital No. 19 of Directive 2008/52/EC which states: "Mediation should not be regarded as a poorer alternative to judicial proceedings in the sense that compliance with agreements resulting from mediation would depend on the good will of the parties. Member States should therefore ensure that the parties to a written agreement resulting from mediation can have the content of their agreement made enforceable".

risk of differential treatment of business users depending on the different national law applicable to that specific dispute[25].

Despite the benefits that mediation could bring, even the practical application of the P2B Regulation by platforms has demonstrated the inefficiency and criticality of the system, which, in fact, has not achieved the desired results.

Looking, in particular, at the implementation and effectiveness of these ODR systems in the market, it is immediately apparent that the tendency has been (and still is) to adhere only formally or, at least, not completely to the prescriptions of the European legislator. Perhaps the most emblematic case is that of Facebook, which has created a "Platform to Business Notice" page[26]. Regarding mediation, the page warns that if the platform fails to resolve the problem through its internal complaint-handling system, the business user may refer the matter to the "Centre for Effective Dispute Resolution" (CEDR) asking for a mediation without, however, directly providing the link to this centre[27]. Indeed, this choice does not facilitate the research for information of the users, with the risk that rather than wasting time searching for all the information concerning the functioning of mediation, they prefer to drop the dispute. The social media provider then gives some brief information on how to initiate the procedure. In particular, the user must notify the other party the fact that a request for a mediation is sent to the CEDR, to which he must forward a copy of the notification. Within the next 14 days, the parties are required to provide the name of a mediator by mutual agreement, and if this is not happening, the CEDR will provide it.

However, looking at the annual report on the application of ODR systems provided by Facebook, we can see how in concrete the use of mediation is very limited and, thus, it seems to be ineffective. In particular, in a period between June 2022 and July 2023, Facebook has received only 751 complaints from users in all over the EU and Instagram even less, i.e., 105. It is quite evident that P2B mediation is not, at least now, a useful tool for the business user who prefer not to establish a dispute.

4. Need for a change in perspective?

The above considerations have brought to light the practical difficulties of alternative dispute resolution systems designed to protect commercial users within the Regulation P2B. In a context of continuous and rapid change such

25 The same recital provides that the MS may refuse to render the mediation agreement enforceable only if: "the content is contrary to its law, including its private international law, or its law does not provide for the enforceability of the content of the specific agreement".

26 To access the Italian 'Platform to Business Notice' page, see: https://m.facebook.com/legal/PlatformtoBusinessNotice.

27 Here the link to the official page of the Centre: https://www.cedr.com/.

as the online environment, it is not sufficient to draw up rules without verifying, once applied in the market, whether they are actually able to protect the fundamental rights of individuals. The P2B Regulation was undoubtedly a long-desired and well-intended intervention, given the need to ensure a fair and transparent online system not only for consumers, but also for professionals who, by now, depends largely on the online intermediation services offered by platforms. However, more than two years after its adoption, it cannot be said that all the critical issues of P2B asymmetrical relationships have been effectively addressed. Alternative dispute resolution tools certainly bring several benefits to the system because they generate a deflective effect of litigation before national courts and, moreover, are capable of better guaranteeing the effectiveness of protection, as they are rapid, freely accessible (almost) by everyone and free (or, at any rate, low-cost). Practical application has shown that platforms have only formally adhered to European requirements, leaving several open questions that not even doctrinal interpretation can resolve.

A major grip came with the recent adoption of the Digital Services Act (henceforth DSA)[28], which imposed stringent transparency and information obligations to fight the diffusion of illegal content online. For this, the regulation has established reporting mechanisms[29] and, in addition to them, effective internal complaint-handling systems to allow users to lodge, electronically and free of charge, a complaint against the decision taken by the platform upon receiving a report[30].

Recipients of such decisions then have the right to turn to any certified out-of-court body to submit to it decisions adopted by the platform as well as complaints that have not been successfully resolved through the in-house complaint-handling instrument[31]. In such a case, platforms are obliged to settle the matter with the body selected by the user in good faith and are then bound by the final decision adopted by that body. Already by reading these provisions, it is possible to notice – albeit limited to online illegal content – the change of perspective of the EU legislator so that, for example, it is envisaged that users can themselves choose which out-of-court body to turn to and the platform, for its part, can do no more than accept the indication and commit itself in good faith. The DSA is, then, a demonstration of how necessary it is to pay

28 European Parliament and Council, Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 December 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) (2022). For a deep analysis, see Rodríguez de las Heras Ballell, T. (2021) 'The background of the Digital Services Act: looking towards a platform economy', *ERA Forum*, vol. 22, pp. 75-86; Chiarella, M. (2022) 'Digital Markets Act (DMA) and Digital Services Act (DSA): New Rules for the EU Digital Environment', *Athens Journal of Law,* vol. 9(1), pp. 33-58.

29 Article 16 of DSA.

30 Article 20 of DSA.

31 Article 21 of DSA.

attention to what is happening in real markets in order to redesign the system to prevent the abuse of digital platforms. There is no doubt that the recipients of information services, both consumers or professionals, are at a distinct disadvantage against the contractual, economic, and technological power of platforms, which end up undermining traditional legal categories. It is still too early to say whether the DSA has achieved its aim of creating a transparent and secure online environment. However, it has introduced additional tools (such as so-called "trusted flaggers")[32] and has imposed stringent transparency and information obligations that do not appear, unlike in the P2B Regulation, to be exclusively formalistic, but translate into substantial protection. This discipline, though, is not aimed at regulating P2B transactions precisely which, therefore, remain regulated almost exclusively by that one regulation which, however, is no longer enough[33].

32 Article 19 of DSA.

33 Rossi, T. (2016) 'Nuovi profili dei rapporti tra imprese nel commercio elettronico', *Informatica e diritto*, vol. XXV, XLII annata, pp. 47-75.

PART IV
Artificial Intelligence, Algorithms and Legal Tech

Chapter XVII
Opening Data of Smart Cities Under the DGA: an Overview of the Challenges Brought About by Data Sharing

by Alessandra Calvi*

Index: 1. New trends in data policies in the European Union. – 2. The Data Governance Act (DGA) in a nutshell. – 3. Data Protection-related challenges of the DGA. – 4. AI fairness-related challenges of the DGA. – 5. Other criticalities of the DGA. – 6. Possible solutions.

1. New trends in data policies in the European Union

Cities across Europe are opening up their datasets and making them available to citizens, other local governments, companies and researchers. It is sufficient to perform a quick search on the internet to come across hundreds of databases containing information about the most diverse subjects, ranging from mobility-related information to energy consumption, from trees planted within a given year to subsidies issued to people with disabilities[1]. Arguably, other than being useful for research purposes[2], making these data held by municipalities available is desirable to promote citizens' scrutiny of local policies. Such openness aligns with a new trend in data policies in the European Union (EU) under the EU Data Strategy[3]. To accelerate the development of artificial intelligence

---

* PhD candidate at the Vrije Universiteit Brussel (VUB) and at CY Cergy Paris Université (CYU), ENSEA, CNRS. At VUB, she is affiliated to the Law, Science, Technology and Society (LSTS) research group and the Brussels laboratory on privacy and data protection impact assessment (d.pia.lab). At CYU, she is affiliated to the Equipes Traitement de l'Information et Systèmes (ETIS) lab, UMR 8051. This work has been funded under the EUTOPIA PhD co-tutelle programme 2021, award number: EUTOPIA-PhD-2021-0000000127 OZRIFTM5. I would like to thank my colleague Pia Groenewolt and the participants of the Information Society Law Center (ISLC) International Conference 2023 for their comments and insights

1 See, for example, Brussels https://opendata.brussels.be/page/home/, Paris https://opendata.paris.fr/pages/home/, Barcelona https://opendata-ajuntament.barcelona.cat/en/.

2 van Eechoud, M. (2022) 'Study on the Open Data Directive, Data Governance and Data Act and Their Possible Impact on Research', *Pubblications Office of the European Union*.

3 European data strategy, available at: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en.

(AI) technologies, data sharing and re-use have gained new momentum. In addition to the Regulation on the free flow of non-personal data[4] and the Open Data Directive (ODD)[5], new rules were adopted, such as the Data Governance Act (DGA)[6] and the Data Act (DA)[7], while others are under discussion, like the AI Regulation (AIR)[8]. Yet, what could be the possible downsides of data sharing, especially for AI development?

Data protection law scholars have already expressed concerns about the difficult reconciliation of data protection and data sharing instruments[9]. By contrast, the possible implications in terms of AI fairness remain largely overlooked.

This contribution contains some preliminary reflections on the challenges brought about by data sharing. It builds on a desk analysis of the DGA, deemed particularly relevant to the smart city context as aimed at strengthening the re-use of public sector information.

2. The Data Governance Act (DGA) in a nutshell

The DGA, which became applicable starting from September 2023 (Article 38 DGA), creates three novel mechanisms to foster data sharing, defined as the "provision of data by a data subject or a data holder[10] to a data user[11] for the

4 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, 28.11.2018.

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

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

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

8 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.

9 De Hert, P. (2023), *Post-GDPR Lawmaking in the Digital Data Society: Mimesis without Integration. Topological Understandings of Twisted Boundary Setting in EU Data Protection Law*, in Deirdre Curtin, D., Catanzariti, M. (eds)(2023) *Data at the Boundaries of European Law*. Oxford University Press; Lindroos-Hovinheimo, S. (2021) 'Data Protection Clashes with Data Sharing : How Will the EU Reconcile Its Two Aims?', *European Law Blog*. Available at: https://www.europeanlawblog.eu/pub/data-protection-clashes-with-data-sharing-how-will-the-eu-reconcile-its-two-aims/release/1.

10 A data holder is "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" (Article 2(8) DGA).

11 By contrast, a data user is "a natural or legal person who has lawful access to certain personal or non-personal data and has the right, including under Regulation (EU) 2016/679 in the case

purpose of the joint or individual use of such data, based on voluntary agreements or Union or national law, directly or through an intermediary" (Article 4(10) DGA)[12].

First, the DGA complements the rules for sharing public sector information encoded in the ODD, by allowing, though not obliging, the re-use of certain data that were not covered by the directive, including personal and protected (on the grounds of commercial or statistical confidentiality, intellectual property, etc.) information held by public sector bodies[13]. Rec. 9 DGA recalls a principle already set in the ODD, namely that public sector bodies and public undertakings need to be encouraged to produce and make available any document (that is, any content or part thereof regardless of its supporting medium (e.g., paper, electronic form, sound, visual or audio-visual recording) following the principle "open by-design and by-default".

Second, it regulates data intermediation services (DIs), namely "those services aiming to establish commercial relationships for the purposes of data sharing between an undetermined number of data subjects and data holders on the one hand and data users on the other […] (Article 2(11) DGA)"[14]. DIs can either support data holders/subjects in granting data users direct access to data (bilateral relationships); or facilitate multilateral relationships and data pools, in which a larger amount of data coming from different data holders/subjects is aggregated into a larger pool[15].

The objective of regulating DIs is twofold. On the one hand, to enhance trust in DIs, whilst contrasting phenomena such as data accumulation and fragmentation[16]. On the other, preventing DIs from abusing their powers and distorting competition[17]. Indeed, despite DIs providers being defined as "neutral" with regard "to the data exchanged between data holders or data subjects

of personal data, to use that data for commercial or non-commercial purposes" (Article 2(9) DGA).

12 Vogel, Y.A. (2022) 'Stretching the Limit, The Functioning of the GDPR's Notion of Consent in the Context of Data Intermediary Services', *European Data Protection Law Review*, vol. 8, p. 238.

13 Namely, the State, regional or local authorities, bodies governed by public law or associations formed by one or more such authorities, or one or more such bodies governed by public law; see also European data strategy, available at: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy_en.

14 Vogel, Y.A. (2022) 'Stretching the Limit, The Functioning of the GDPR's Notion of Consent in the Context of Data Intermediary Services', *European Data Protection Law Review*, vol. 8, p. 238.

15 Ibid.

16 The former refers to the accumulation of data within only few big players. The latter to the inability of smaller entities to locate and agglomerate data from disparate sources. von Ditfurth, L., Lienemann, G. (2022) 'The Data Governance Act: – Promoting or Restricting Data Intermediaries?' *Competition and Regulation in Network Industries*, vol. 23.

17 Ibid.

and data users" (rec. 33)[18], risks in terms of anti-competitive practices due to the reliance on intermediaries have been documented[19]. Accordingly, DIs are for example prevented from using data provided by data holders for their own commercial purposes (article 12(a) DGA)[20]. When offered by organisational structures composed of data subjects, one-person undertakings or SMEs aimed at supporting their members in the exercise of their rights with respect to certain data, as well as exchanging views on data processing purposes and identifying conditions that best represent the interests of their members, data intermediation services become "services of data cooperatives" (article 2(15) DGA). Such DIs, including data cooperatives, need to comply with certain conditions (article 12 DGA) and undergo a notification procedure (article 11 DGA)

Third, the DGA introduces the concept of data altruism, which is the:

> voluntary sharing of data on the basis of the consent of data subjects to process personal data pertaining to them, or permissions of data holders to allow the use of their non-personal data without seeking or receiving a reward […] for objectives of general interest [e.g., healthcare, fight against climate change, improving mobility, production and dissemination of statistics] as provided for in national law (article 2(16) DGA).

Data altruism organisations need to comply with some substantive requirements (e.g., operate on a not-for-profit basis and be legally independent of any entity that operates on a for-profit basis) as well as transparency obligations, and may undergo a registration process (Chapter IV DGA).

All these developments will affect data sharing in smart cities, by facilitating information exchanges among citizens, local authorities and tech providers to achieve the "common good"[21]. In principle, data altruism organisations and data cooperatives could enhance the capacity of data subjects to control their personal information, which is usually challenged due to *inter alia* ubiquitous data processing and information asymmetries existing in the smart city context[22]. They could also foster the inclusion of data subjects in the smart city discourse and possibly promote the adoption of bottom-up smart city initiatives. These developments are desirable considering that, currently, despite being major producers of

18 Vogel, Y.A. (2022) 'Stretching the Limit, The Functioning of the GDPR's Notion of Consent in the Context of Data Intermediary Services', *European Data Protection Law Review*, vol. 8, p. 238.

19 von Ditfurth, L., Lienemann, G. (2022) 'The Data Governance Act: – Promoting or Restricting Data Intermediaries?', *Competition and Regulation in Network Industries*, vol. 23.

20 Ibid.

21 See https://living-in.eu/news/webinar-living-ineu-legal-subgroup-data-governance-act-proposal-and-its-impact-european-cities.

22 Calvi, A. (2021) 'A Step Towards Dystopia? How the COVID-19 Pandemics Exacerbated the Data Protection Challenges Raised by Smart Cities', *Ethics, Health Data, and Bio-Citizenship*, vol. 22, p. 205.

information, citizens (data subjects, in data protection jargon) are usually left at the margin of the smart city debate, especially when lacking representation among the more powerful smart city actors, such as tech providers and municipalities[23]. However, as it will be illustrated hereafter, some criticalities remain.

3. Data Protection-related challenges of the DGA

Admittedly, the data protection literature on the DGA is still quite scarce. Most works covered the possible inconsistencies between the General Data Protection Regulation (GDPR) and the – back then – DGA proposal and are currently outdated since the EU co-legislators took on board several recommendations provided by the European Data Protection Supervisor (EDPS), the European Data Protection Board (EDPB) and academics to improve the DGA draft.

As regards the protection of personal data, article 1(3) DGA states that the DGA is without prejudice to the GDPR, meaning that in case of conflicts, the latter shall prevail. Besides, article 2 DGA clarifies that notions such as personal data, consent, data subject and processing when referring to personal data need to be interpreted as in the GDPR. This represents an improvement compared to the original DGA draft, which was criticised by data protection regulators and some commentators due to the insufficient coordination between the two legal instruments[24]. Other clarifications in terms of coordination between the two instruments are contained in the Recitals, too. For instance, Rec. 7 DGA expressly calls for Member States to support public sector bodies in applying "state-of-the-art privacy-preserving methods that could contribute to a more privacy-friendly processing of data", such as anonymisation, differential privacy, generalisation, suppression and randomisation, synthetic data, other than referring to the importance of Data Protection Impact Assessments (DPIAs). Rec. 4 states that the DGA should be without prejudice to the GDPR including where personal and non-personal data in a data set are inextricably linked, thus clarifying the rules applicable to mixed datasets, namely those consisting of both personal and non-personal data and that according to the EU Commission will constitute the majority of datasets in the data economy[25].

23 Calvi, A. (2022) 'Gender, Data Protection & the Smart City: Exploring the Role of DPIA in Achieving Equality Goals', *European Journal of Spatial Development (EJSD)*, vol. 19, p. 24.

24 EDPB-EDPS, Joint Opinion 03/2021 on the Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act), 2021; Gellert, R., Graef, I. (2019) 'The European Commission's Proposed Data Governance Act: Some Initial Reflections on the Increasingly Complex EU Regulatory Puzzle of Stimulating Data Sharing', *TILEC Discussion Paper*, vol. 6.

25 Ibid.

Yet, some uncertainties remain. As anticipated, the coordination of data-sharing-focused instruments with the GDPR raises a number of questions[26]. The DGA covers both personal and non-personal data at the same time, and the definition of data given therein is extremely broad, as "any digital representation of acts, facts or information and any compilation of such acts, facts or information, including in the form of sound, visual or audio-visual recording" is considered data (Article 2(1) DGA)[27]. Since the DGA does not always specify whether its provisions apply to personal or non-personal data or both and given that the distinction between personal and non-personal data fades in the context of big data, this could lead to uncertainties as to the rules concretely applicable in a given context[28]. Then, although it was clarified that the notion of "data holder" does not encompass "data subject", the relationship thereof with the notions of (joint) controllers, processors and recipients remains unclear [29]. Similarly, the notion of "data user"[30]. Other scholars emphasised how, regardless of the letter of the law, applying the notion of consent under the GDPR proves to be impossible in the context of multilateral data relationships (those involving an undetermined number of data users and an indefinite number of data holders/subjects) facilitated by DIs. Compliance with the strict consent requirements (namely, unambiguous, informed, freely given and specific) under the GDPR indeed could not be granted[31].

4. AI Fairness-related challenges of the DGA

Unlike data protection issues, the AI fairness implications arising from data sharing have been largely overlooked in the DGA. This is an important gap to address considering that data sharing aims, among others, to promote AI development. The definition fairness is domain specific, which makes the

26 De Hert, P. (2023) *Post-GDPR Lawmaking in the Digital Data Society: Mimesis without Integration. Topological Understandings of Twisted Boundary Setting in EU Data Protection Law*, in Deirdre Curtin, D., Catanzariti, M. (2023)(eds) *Data at the Boundaries of European Law*. Oxford University Press; Lindroos-Hovinheimo, S. (2021) 'Data Protection Clashes with Data Sharing : How Will the EU Reconcile Its Two Aims?', *European Law Blog*. Available at: https://www.europeanlawblog.eu/pub/data-protection-clashes-with-data-sharing-how-will-the-eu-reconcile-its-two-aims/release/1.

27 Baloup, J., et al. (2021) 'White Paper on the Data Governance Act', *CiTiP Working Paper 2021*. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3872703.

28 European Data Protection Board, Statement on the Digital Services Package and Data Strategy Adopted on 18 November 2021.

29 EDPB-EDPS, Joint Opinion 03/2021 on the Proposal for a Regulation of the European Parliament and of the Council on European Data Governance (Data Governance Act), 2021; Baloup, J., et al. (2021) 'White Paper on the Data Governance Act', *CiTiP Working Paper 2021*.

30 Ibid.

31 Vogel, Y.A. (2022) 'Stretching the Limit, The Functioning of the GDPR's Notion of Consent in the Context of Data Intermediary Services', *European Data Protection Law Review*, vol. 8.

operationalisation and contextualisation of this concept in the context of AI a difficult exercise[32]. For legal scholars, fairness is a general concept strictly related to non-discrimination[33], that appears in several areas of law, ranging from contracts to criminal justice to competition and consumer law. The DGA mentions the need to safeguard "fair competition" and mandates DIs to offer access to their services under "fair, transparent and non-discriminatory conditions" for data subjects, data holders and data users (Article 12(f) DGA). Then, Rec. 2 refers to the so-called "FAIR data principles" namely that data should be made Findable, Accessible, Interoperable and Re-usable, while ensuring a high level of cybersecurity. The FAIR data principles were elaborated as best practices in the context of data management, by putting specific emphasis on enhancing the ability of machines to automatically find and use the data, in addition to supporting their reuse by individuals[34]. However, despite not being expressly recalled by the DGA in this sense, fairness is also a general principle of personal data processing (Article 5(1)(a) GDPR), linked to the protection of the reasonable expectations data subjects[35], to safeguard their autonomy in determining the use of their personal data; not discriminating against them; not exploiting their needs and vulnerabilities; addressing power imbalances[36].

Conversely, for computer scientists, fairness is a mathematical property for algorithms[37]. Multiple (and often incompatible) mathematical definitions of fairness exist, corresponding to multiple fairness metrics[38]. Fairness metrics

32 Calvi, A., Kotzinos, D. (2023) 'Enhancing AI Fairness through Impact Assessment in the European Union: A Legal and Computer Science Perspective', *ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*.

33 Wachter, S., Mittelstadt, B., Russell, C. (2021) 'Why Fairness Cannot Be Automated: Bridging the Gap between EU Non-Discrimination Law and AI', *Computer Law and Security Review*, vol. 41. Available at: https://doi.org/10.1016/j.clsr.2021.105567; Zuiderveen Borgesius, F., (2018) *Discrimination, Artificial Intelligence and Algorithmic Decision Making*. Available at: https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73; Binns, R. (2018) 'Fairness in Machine Learning: Lessons from Political Philosophy', *Machine Learning Research (Conference on Fairness, Accountability, and Transparency 2018)*. Available at: http://arxiv.org/abs/1712.03586; Barocas, S., Selbst, A.D. (2016) 'Big Data's Disparate Impact', *California Law Review*, vol. 104, p. 671.discrimination, and fairness in artificial intelligence (AI)

34 Wilkinson, M.D., et al. (2016) 'Comment: The FAIR Guiding Principles for Scientific Data Management and Stewardshi'. *Scientific Data*, vol. 3.

35 European Data Protection Board, 'Guidelines 2/2019 on the Processing of Personal Data under Article 6(1)(b) GDPR in the Context of the Provision of Online Services to Data Subjects'.

36 European Data Protection Board, 'Guidelines 4/2019 on Article 25 Data Protection by Design and by Default v.2.0'.

37 Green, B., Hu, L. (2018) 'The Myth in the Methodology: Towards a Recontextualization of Fairness in Machine Learning', *35th International Conference on Machine Learning*.

38 Ibid; Makhlouf, K., Zhioua, S., Palamidessi, C. (2020) 'Machine Learning Fairness Notions: Bridging the Gap with Real-World Applications', *Computer Science*. Available at: https://arxiv.org/abs/2006.16745.

aim at addressing the problem of bias, occurring when automated systems systematically and unfairly discriminate against certain individuals or groups of individuals in favour of others (e.g., by denying an opportunity/good or assigning undesirable outcomes based on inappropriate grounds)[39].

Regrettably, in the DGA, there is no mention of any quality criteria that data supposed to be shared and used for AI training need to comply with to prevent bias[40]. This is concerning considering that bias largely depends on data quality. Where the data used to train an algorithm are either incomplete or unrepresentative (data gap) or conversely disproportionately abundant (data overload), this could lead to biased results[41].

Someone may argue that these issues would be better dealt with in the AIR, which indeed fixes some data governance and management requirements for datasets used for training, validation and testing purposes (see e.g., Article 10 AIR). And that bias can still be treated at later stages of the AI life cycle. However, as the AIR is still under discussion, currently there is a legal vacuum. Then, the safeguards in the AIR are framed rather in terms of best-effort obligations and apply exclusively to high-risk AI systems. Furthermore, addressing bias later may result in higher costs[42]. Then, the lack of coordination between the DGA and the AIR, and between the duties existing upon public sector bodies, DIs and data altruism organisations, on the one hand, and developers, on the other, could result in inconsistencies.

Even the prominence given to Privacy Enhancing Technologies (PETs) in the DGA is problematic in terms of AI fairness. Rec. 7 DGA advocates for the use of "state-of-the-art privacy-preserving methods that could contribute to a more privacy-friendly processing of data", such as anonymisation, differential privacy, generalisation, suppression and randomisation, synthetic data. Public sector bodies are particularly encouraged to adopt them to promote data sharing. Whereas these techniques are indeed important to ensure privacy, some of them may jeopardise bias detection. Regrettably, the DGA seems to completely ignore the trade-offs between the application of PETs and AI fairness, despite the last years computer science scholarship has been increasingly raising

39 Friedman, B., Nissenbaum, H. (1996) 'Bias in Computer Systems', *ACM Transactions on Information Systems*, vol. 14, p. 330.

40 Calvi, A. (2023) 'Exploring the Synergies between Non-Discrimination and Data Protection: What Role for EU Data Protection Law to Address Intersectional Discrimination?', *European Journal of Law and Technology*, vol. 14(2), p.1.

41 Turner Lee, N., Resnick, P., Barton, G. (2021) 'Algorithmic Bias Detection and Mitigation: Best Practices and Policies to Reduce Consumer Harms', *Brookings*. Available at: https://www.brookings.edu/research/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/.

42 Calvi, A., Kotzinos, D. (2023) 'Enhancing AI Fairness through Impact Assessment in the European Union: A Legal and Computer Science Perspective', *ACM Conference on Fairness, Accountability, and Transparency (FAccT '23)*.

awareness about this issue[43]. In particular, differential privacy is a PET that ensures privacy by adding random noise to a dataset. In other words, it promotes the protection of privacy by introducing some error to a dataset, thus affecting the accuracy thereof. This technique makes it possible to gather statistically significant insights whilst preventing the re-identification of individuals, since, by looking at the output (namely, the new differentially private dataset), it would not be possible to tell whether any individual's data were included in the original dataset or not[44]. Whether in principle differential privacy should work for any individual for any dataset, studies demonstrated that, when strict privacy settings apply, and even when privacy mechanisms add equivalent noise to independent populations, significant disparities in the outcomes of the algorithm may nevertheless occur, and thus some populations may be more affected than others[45]. Other studies highlighted how differential privacy may fail to provide "fair privacy protection", being possible that the likelihood and/or the cost of a privacy failure affect users differently, depending on protected characteristics[46].

5. Other criticalities of the DGA

The nature of data cooperatives and data altruism organisations remains contested. Despite data cooperatives portraying themselves as drivers of collaborative approaches to the production, use, monetisation of and access to data,

43 Agarwal, S. (2019) 'Trade-Offs between Fairness and Interpretability in Machine Learning', *IJCAI Workshop on AI for Social Good*; Cumming, R., et al. (2019) 'On the Compatibility of Privacy and Fairness', *ACM UMAP 2019 Adjunct – Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*; Kiss, M. (2021) *Demographic Outlook for the European Union 2020*. Available at: http://www.europarl.europa.eu/thinktank; Ekstrand, M.D., Joshaghani, R., Mehrpouyan, H. (2018) 'Privacy for All: Ensuring Fair and Equitable Privacy Protections', *Proceedings of Machine Learning Research*. Available at: https://proceedings.mlr.press/v81/ekstrand18a.html; Pujol, D., et al. (2020) 'Fair Decision Making Using Privacy-Protected Data', *FAT* 2020 – (Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*). Available at:: https://arxiv.org/abs/1905.12744.

44 Calvi, A. (2021) 'Differentially Private Algorithms', *Privacy Laws & Business*, p. 11; Pujol, D., et al. (2020) 'Fair Decision Making Using Privacy-Protected Data', *FAT* 2020 – (Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*); Ekstrand, M.D., Joshaghani, R., Mehrpouyan, H. (2018) 'Privacy for All: Ensuring Fair and Equitable Privacy Protections', *Proceedings of Machine Learning Research*. Available at: https://proceedings.mlr.press/v81/ekstrand18a.html.

45 Pujol, D., et al. (2020) 'Fair Decision Making Using Privacy-Protected Data', *FAT* 2020 – (Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*).we initiate a first-of-its-kind study into the impact of formally private mechanisms (based on differential privacy

46 Ekstrand, M.D., Joshaghani, R., Mehrpouyan, H. (2018) 'Privacy for All: Ensuring Fair and Equitable Privacy Protections', *Proceedings of Machine Learning Research*. Available at: https://proceedings.mlr.press/v81/ekstrand18a.html.

they still build upon a neoliberal understanding thereof[47]. Rec. 31 DGA states that their primary purpose is to:

> strengthen the position of individuals in making informed choices before consenting to data use […] - and give - better choices to the individual members of the group or potentially finding solutions to conflicting positions of individual members of a group on how data can be used where such data relates to several data subjects within that group.

Thus, still quite an individualistic understanding of the right to data protection aimed rather at avoiding clashes among multiple individual interests than moving towards a collective understanding of data protection[48].

As regards data altruism organisations, it is quite likely that factors such as gender, race, abilities, age, digital literacy will affect the trust of data subjects as well as their decisions on whether to rely on them. Thus, data altruism organisations can either become instruments of empowerment or conversely a vessel to obfuscate the concerns of data subjects belonging to vulnerable and marginalised groups. Accordingly, it is also possible that data shared by data altruism organisation carry some bias as some groups may be over or under share compared to others. Furthermore, it was noted how the very same idea of data altruism assumes the desirability of data sharing and pushes data subjects toward that. Instead, the effects thereof may be perverse[49]. Whereas a data subject may not suffer individual harm by sharing information (e.g., when accepting to share their biometrics to train facial recognition), her choice could have nevertheless collective impacts, such as the further marginalisation of certain groups (e.g., if such facial recognition system is then used to push away migrants)[50].

Finally, the DGA lacks a definition of general interest, which is conversely a condition for setting up data altruism schemes, leaving it largely to national law. As already observed about the notion of (substantial) public interest[51], even the concept of general interest is so volatile that could be used to justify whatever policy, undermining the consent of data subjects. For instance, the prevention of crime could be considered an objective of general interest, but the ways to achieve it are multiple, ranging from the implementation of surveillance technologies to the reform of the criminal system and the introduction of welfare

47 Bietti, E., Etxeberria, A., Mannan, M., Wong, J. (2021) *Data Cooperatives in Europe: A Legal and Empirical Investigation*. Available at: https://cyber.harvard.edu/sites/default/files/2022-02/Data_Cooperatives_Europe-group2.pdf.

48 Ibid.

49 Garfield, B. (2021) 'What We Do with Data: A Performative Critique of Data "Collection"', *Internet Policy Review*, vol. 10.

50 Ibid.

51 Christofi, A., Wauters, E., Valcke, P. (2021) 'Smart Cities, Data Protection and the Public Interest Conundrum: What Legal Basis for Smart City Processing?', *European Journal of Law and Technology*, vol. 12.

measures. A data subject may agree with the objective of general interest but not with the means proposed to achieve said objective.

6. Possible solutions

To address the challenges illustrated above, it is necessary to act at multiple levels. Assuming the impracticability of amending such a recent regulation, the first step is to raise awareness among public sector bodies, DIs, data altruism organisations and data subjects, as well as regulators, about the challenges brought about by data sharing. In particular, about the interplays and trade-offs between PETs and AI fairness, which, contrary to data protection concerns, appear totally overlooked. It must be clear that not collecting or aggregating demographic information may prevent the investigation of discrimination issues in databases[52] and that PETs may have different impacts on people depending on their personal – even protected – characteristics[53]. Addressing these concerns is crucial in the context of smart cities. They are enormous collectors, users and sharers of (personal) data and datasets. Considering how much local authorities rely on data to inform their political activities, biased information could exacerbate discrimination and oppression dynamics already existing in the urban context[54]. Some technical solutions to address these issues have been investigated and proposed, such as customising privacy mechanisms depending on policy goals[55]. Yet, further research is needed. And guidance from EU regulators about how to balance between the two would be desirable. In the meantime, a possible means to make trade-offs between PETs and fairness issues emerge is through DPIAs, whose importance "to more safety in the use and re-use of personal data" is reiterated in Rec. 7. Yet, some uncertainties remain as to how the obligation to carry out a DPIA (which covers exclusively personal data processing) will be distributed among the different actors involved in data sharing under the DGA.

A clearer definition of "objectives of general interest" is desirable too, as this concept is extremely elusive. Admittedly, transparency requirements for

52 Ekstrand, M.D., Joshaghani, R., Mehrpouyan, H. (2018) 'Privacy for All: Ensuring Fair and Equitable Privacy Protections', *Proceedings of Machine Learning Research*. Available at: https://proceedings.mlr.press/v81/ekstrand18a.

53 Pujol, D., et al. (2020) 'Fair Decision Making Using Privacy-Protected Data', *FAT* 2020 – (Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency)*; Cumming, R., et al. (2019) 'On the Compatibility of Privacy and Fairness', *ACM UMAP 2019 Adjunct - Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization*.we initiate a first-of-its-kind study into the impact of formally private mechanisms (based on differential privacy

54 Calvi, A. (2022) 'Gender, Data Protection & the Smart City: Exploring the Role of DPIA in Achieving Equality Goals', *European Journal of Spatial Development (EJSD)*, vol. 19.

55 Pujol, D., et al. (2020) 'Fair Decision Making Using Privacy-Protected Data', *FAT* 2020 – (Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency)*.

data altruism organisations aimed *inter alia* at clarifying the "objectives of general interests" grounding data processing already exist (see Article 20 and 21 DGA), but it is unclear to what extent the definition needs to be precise. Clarifying them as much as possible is necessary to ensure that data subjects make informed choices and agree not only with the objectives broadly speaking but also with the concrete means proposed to achieve them. Then, there is no reference, for instance, to the need to inform data subjects about the possible societal consequences of data sharing, only about possible misuse of data. These clarifications and further specifications may be provided through the Rulebook mentioned in Article 22 DGA, which is supposed to be adopted by the Commission via implementing acts and is aimed *inter alia* at clarifying information duties existing on data altruism organisations towards data subjects and data holders. Possibly, some clarifications as to how to comprehensively provide information to data subjects could be made even in the European Data Altruism Consent form (Article 25 DGA), which is also supposed to be adopted by the Commission via implementing acts as well.

Chapter XVIII
The Algorithm in Administrative Decisions: Risks and Opportunities

by Susanna Viggiani*

INDEX: 1. Artificial intelligence in Public Administration. – 2. Automated decisions: the essential principles for the protection of data subjects' rights. – 3. The risks linked with the algorithmic decision.

1. Artificial intelligence in Public Administration

Today, the digitisation process of public administration is considered an unstoppable and constantly evolving process. The introduction of new technological tools would allow for changes in both organisation and administrative activity. Already in 1979, the Minister for the Public Function, Massimo Severo Giannini, in his «Report on the main problems of State administration», noted that:

> computers, originally used as devices for the simple registration of complex data, have become devices for assessment and verification, for calculation, for articipation in procedural stages of investigation, and finally for decision-maing, so that computer systems no longer serve administrations for internal management purposes, but are used specifically for administration, i.e. they are increasingly projected outwards[1].

Indeed, the process of technological innovation - completed to date - has enabled the transition from an administration that relied only on paper tools to an administration able to manage digital files and to deliver services digitally (think of digital identification services – SPID – or the PagoPA online payment systems). Consequently, today we are in that phase called by administrative law experts "the fourth phase of Public Administration" because it is characterised, like Industry 4.0, by a high rate of automation and interconnection and in which the processing, storage and transformation of data through the use of Artificial Intelligence (henceforth A.I.) systems plays a central role. A.I. systems – which

---

* Legal and Privacy specialist, specialized in Public Administration studies (SPISA).

1 Giannini, M.S. (1979) *Il rapporto sui principali problemi dell'amministrazione dello Stato*. Ministro per la funzione pubblica.

many scientists call Alternative Intelligence Systems – are able to perform a set of functions that, in the past, were only attributed to humans. So,

> if the process is qualified as intelligent when it is performed by a human being, then it can also be qualified as intelligent if it is performed by a machine. So Artificial Intelligence can be defined the science of making computers do things that would require intelligence when they are done by humans[2].

One of the inherent problems of A.I. depends directly on the specific nature of such systems, particularly when they employ complex algorithmic compositions. More complex the algorithm, better its performance capability, but bigger the risks connected to its use. There are, in fact, different algorithmic categories: there are conditional or deterministic algorithms and predictive algorithms. The first allow predetermined rules to be applied when certain conditions come true, as happens in the restricted activity of the P.A. While the second used in machine learning technologies, are composed of two essential elements: the source code and the mathematical model. In short, the machine learning system generates parameters, based on a wealth of information, consisting of datasets (input), that was provided to it during the training phase. This algorithm is then able to examine the degree of correspondence between the result provided and the mathematical model and finally come to a conclusion[3]. The same procedure is used by the administration in the exercise of discretionary administrative activity, according to which it can only take a specific decision following the assessment and balancing of a primary public interest with secondary public, private or collective interests.

Consequently, the use of complex algorithms in administrative activities and processes allows for a higher level of discretion, but this could put the traditional categories of administrative law to the test, without an appropriate regulatory measure[4]. From the point of view of administrative law, the introduction of the CAD – Digital Administration Code – and, in particular, the structural and managerial reorganisation of the P.A. with the use of new technologies introduced by Article 15 would seem to be aimed at combating the inefficiencies of the administration, also through the adoption of organisational models capable of enforcing the use of information technology within the administrations themselves and in relations with other P.A. and private parties. Administrative jurisprudence, after some initial doubts[5] between automated procedures and

2 Finocchiaro, G. (2020) 'Intelligenza artificiale e responsabilità', *Contr. e impr.*, vol. 2, p. 713.

3 Cons. St., Sez. VI, sent. n. 2270/2019, Canalini, V. (2019) 'L'algoritmo come "atto amministrativo informatico" e il sindacato del giudice', *Giornale di diritto amministrativo*, vol. 6, pp. 781-787.

4 Orsoni, G., D'Orlando, E. (2019) 'Nuove prospettive dell'Amministrazione digitale: Open data e algoritmi', *Istituzioni del Federalismo*, vol. 3, pp. 593.

5 Cons. St., sez. VI, n. 2270/2019.

L. 241/1990 – the Law on Administrative Proceedings – has accepted these artificial intelligence tools, on condition that they respect certain principles of algorithmic legality[6]. According to these principles, the algorithm has legal value and must comply with the principles of administrative activity; it cannot be discretionary, in the sense that discretion must be adopted only at the time the digital tool is developed; it must respond to purposes predetermined by law, in order to comply with the principle of legality; it must be monitored and updated over time and must be subject to review by the judge, who must verify the propriety of the automated process in all its components, according to the paradigm of due process[7].

The European Union, too, recently adopted a legislative act dealing with the complex phenomenon of artificial intelligence, in order to give the Member States a common framework of rules: this is the Artificial Intelligence Regulation, approved in 2023 and still to be published. This contains a risk-based approach, distinguishing between four types of risk: unacceptable, high, minimal and limited. Depending on the risk category, a different regulatory system applies. The act is part of a set of regulations aimed at regulating the digital dimension of human life, such as the Data Governance Act (EU Regulation 2022/868), the Digital Services Act (EU Regulation 2022/2065), the Digital Markets Act (EU Regulation 2022/1925), the Data Act, proposed on 23 February 2022. Moreover, A.I. is powered by data of human beings, which need specific protections from the institutions. Among the recognised protections, EU Reg. 679/ 2016 (henceforth GDPR) and Legislative Decree 196/2003 as modified by Legislative Decree 101/2018 (henceforth Privacy Code) extend the discipline put in place to protect privacy to all parties involved. In fact, in order to guarantee the right to know the logic underlying the administrative decision, pursuant to Article 13 co. 2 lett. f) of the GDPR, the administrative judge ruled that the algorithm must be transparent, i.e. it must recognise the data subject's right to dispose of the source code because it is an essential part of the algorithm. It follows, that also the mechanism through which the decision elaborated by the algorithm is realised, in order to be effective, must be tested under the parameter of legality and the principle of transparency[8].

6 Carloni, E. (2020) 'I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo', *Diritto amministrativo*, vol. 2, pp. 273-304.

7 Linee guida del gruppo di Alti esperti in materia di I.A. Available at: https://ec.europa.eu/digital-single-market/en/high-level-expert-group-artificial-intelligence; Risoluzione del Parlamento europeo del 12 febbraio 2019 su una politica industriale europea globale in materia di robotica e intelligenza artificiale (2018/2088). Available at: http://www.europarl.europa.eu/doceo/document/TA-8-2019-0081_IT.html.

8 Manganaro, F. (2012) *L'evoluzione del principio di trasparenza,* in Scoca, F.G. (2012) *Studi in memoria di Roberto Marrama,* Editoriale Scientifica, Napoli, p. 639.

2. Automated decisions: the essential principles for the protection of data subjects' rights

The Sixth Section of the Council of State in several sentences[9] has encouraged the use of A.I. systems on the basis of the principles of good performance and impartiality of administrative action, Article 97 of the Constitution, regarded as the main values to which the public administration must aspire. In addition, even doctrine and jurisprudence have highlighted the advantages that the digital processing act would present because it could reduce the risks of non-compliance with the law and the prevention of corruption.

If administrative decisions are based on algorithms, this would have the advantage of being able to solve complex issues with rational, efficient and potentially neutral tools. However, larger margin of discretion, recognized in the head of the Administration, the higher will be the difficulty to offer objective and calculable the decision of the machine, the operation of the algorithm and the logical-juridical iter followed by the system.

For these reasons, the European Parliament Resolution of 12 February 2019 – on a global European industrial policy on robotics and artificial intelligence – highlights the need for algorithmic decisions to be intelligible and knowable by the subject involved in this process. In order to ensure proper compliance with the legislation, transparency towards the user should be ensured, and on the other hand, accountability and accountability[10] of those who have developed and use artificial intelligence solutions[11]. In fact, article 22 of the GDPR introduces the right to the data subject not to be subjected – and, consequently, the relative right to oppose – to a decision based just on automated data processing – including profiling – which has legal effects affecting him or which have a significant effect on him[12].

As a result, it would not be allowed a final action that is limited to transposing the result provided by the algorithm. This would exclude the hypothesis of fully automated procedures. The administration must always give a reason for the measure. However, this is very difficult when using machine learning systems[13].

---

9 Consiglio di Stato, Sez. VI, feb. 2020, n. 881; Consiglio di Stato, Sez. VI, dec. 2019, n. 8472-8473-8474; Consiglio di Stato, Sez. VI, apr. 2019, n. 2270.

10 Article 24 GDPR: "Taking into account the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for the rights and freedoms of natural persons, the controller shall implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing is performed in accordance with this Regulation. Those measures shall be reviewed and updated where necessary".

11 D'Aloia, A. (2019) 'Il diritto verso "il mondo nuovo". Le sfide dell'Intelligenza Artificiale', *BioLaw Journal*, vol. 1, p. 19.

12 Recital 71.

13 Tar Campania, Sez. III, n. 7003/2022; Cons. St., n. 7891/2021.

The principle of European law on the non-exclusivity of algorithmic decision is therefore fundamental. This principle states that the decision-making process must always be accompanied by a human contribution. This has the task of checking, confirming or denying the automatic decision. The main task of the human operator is always to assure the rights of transparency, impartiality and non-discrimination to those involved in such decisions. According to the judges, the algorithm and artificial intelligence can only be used in the public area if they respect the general principles of administrative activity, such as transparency, reasonableness and proportionality. These systems must be visible and subject to analysis, judgment and evaluation by the judge.

3. The risks linked with the algorithmic decision

One of the most obvious risks in the matter of algorithmic decision-making is certainly that relating to the opacity of the decision-making process followed by public administrations[14]. The question for the administrative courts essentially concerns the legality of the use of computer technology to determine the content of administrative acts. The use of such technologies of I.A. would offer benefits to human activity in terms of utility and effectiveness, but what worries, it is the change of the subject that makes the decision because it would go from a decision taken by man to an algorithmic decision based on mathematical formulas.

The main problems related to the use of algorithms in the activities of the P.A. are three: the explicability and autonomy of action; the control of the result that is how to ensure the legitimacy of the administrative act and finally, the principle of non-discrimination algorithm. For what concerns the explicability of the algorithm, in the administrative activity the choices must always be made on the basis of a power provided by a given norm. It is always necessary to know who must exercise the power and how it must be exercised. The use of machine learning algorithms, on the other hand, creates problems related to the explicability of the procedure followed to make the decision because this is taken through learning processes that are difficult to explain and understand.

With reference to the level of autonomy and control of the result, the information that the Administration acquires to decide are taken through technological platforms. However, Italian case-law has stated that "the computer procedures, even if they have their greater degree of precision and are close to perfection, can never change and replace the cognitive activity of thought and judgment that only a natural person official is able to carry out"[15]. This

14 Coglianese, C., Lehr, D. (2017) 'Regulating by Robot. Administrative Decision Making in the Machine-Learning Era', *Georgetown Law Journal*, vol. 105.

15 Tar Lazio, sez. 3-bis, oct. 2018, n. 9224.

meant that the result obtained by the algorithm must necessarily be controlled by a natural person official. The natural person must test that the final decision complies with all the principles of good administration.

According to non-discrimination principle it is necessary that the algorithm is not inherently discriminatory: the data controller must adopt appropriate organisational measures in order to guarantee and prevent discriminatory effects. Otherwise, there is a violation of the principle of non-discrimination laid down in the Charter of Fundamental Rights of the European Union, the European Convention on Human Rights and Italian Constitution, which protects the principle of equality. Artificial intelligence algorithms use data to learn, to know and to bring out correlations[16]. Who rules the artificial intelligence solutions, to ensure quality and compliance[17] with the requirements must ensure a monitoring system and must report any serious incident or malfunction that may be a breach of obligations that must protect the fundamental rights of human beings. The European regulation, in fact, provides for supervisory and control mechanisms and a system of sanctions in case of violation of the provisions[18].

As well as the risk of opacity and invisibility of algorithmic administration, the doctrine highlighted the risk of reaching discriminatory[19] outcomes mainly due to: the identification and selection of "target variables" and "class labels"; the collection and selection of data (data training); the selection of the characteristic inserted in the model ("feature selection"); the choice of the "proxy"; the intentional discrimination, that is consciously placed in being from the programmer[20]. The violation of the principle of non-discrimination[21] could be due to the poor quality of the data acquired by the software. That is why specific moments of control of the procedural outcomes reached by the machine are necessary, in order to verify that the rules of operation of the automated procedure – defined *ex ante* – are respected before the adoption of the administrative measure. The administrative decision can be supported by an algorithm, providing predictions, but can never completely replace it.

16 Amato Mangiameli, A.C. (2022) *Intelligenza artificiale, big data e nuovi diritti*, in Abba, L., Lazzaroni, A., Pietrangelo M. (2022) *La Internet governance e le sfide della trasformazione digitale,* Rivista Italiana di Informatica e Diritto, p. 97.

17 Finocchiaro, G. (2020) *Diritto di Internet,* Zanichelli Editore, Bologna, III ed., p. 188.

18 Article 14 Artificial Intelligence Act, COM (2021) 206 final – 21 april 2021.

19 Costantino, F. (2019) 'Rischi e opportunità del ricorso delle amministrazioni alle predizioni dei big data', *Diritto pubblico*, vol. 1, p. 56; Cavallaro, M.C., Smorto, G. (2019) 'Decisione pubblica e responsabilità dell'amministrazione nella società dell'algoritmo', *Federalismi.it*, vol. 16, p. 19.

20 Resta, G. (2019) 'Governare l'innovazione tecnologica: decisioni algoritmiche, diritti digitali e principio di uguaglianza', *Politica del diritto*, p. 214.

21 Consider the proposal under discussion in the institutions of the European Union by which they would like to introduce a ban on the use of facial recognition AI techniques because of their high rate of discrimination.

For that reason, an unreasonable and undisciplined use of technology could pose a serious risk to the fundamental rights and freedoms of data subjects. Monitoring is essential, in order to prevent and combat the spread of intrusion, manipulation and alteration by the increasingly dangerous cybercriminals. In this sense, fundamental is the role played lately by AgID and National Agency of Cybersecurity, which have planned a programme of initiatives. This provides to raise awareness the public administrations on issues related to cybersecurity, especially when these make use of artificial intelligence systems[22].

22 See https://www.agid.gov.it/sites/default/files/repository_files/pianotriennaleinformatica-pa2021-2023.pdf.

Chapter XIX
The AI Act Proposal: a New Right to Technical Interpretability?

by Chiara Gallese*

INDEX: 1. Introduction. – 2. The concepts of "technical interpretability" and "explainability". – 3. The transparency principle and the "right to explanation". – 4. Articles 13 and 14 of the AI Act Proposal. – 5. The "right to technical interpretability" as a fundamental right.

1. Introduction

The debate about the concept of "right to explanation" is the subject of a wealth of literature. It has focused in the legal scholarship on article 22 GDPR and in the technical scholarship on explaining the output of a certain model. The purpose of this paper is to investigate if the new provisions introduced by the proposal for a Regulation laying down harmonised rules on artificial intelligence (AI Act), in combination with Convention 108+ and GDPR, are enough to indicate the existence of a "right to technical explainability" and, if not, whether the EU legal system should include it in its current legislation.

Before examining the legal aspects of "interpretability", it is necessary to provide clear definitions and to differentiate it from "explainability" (Section 2). From a legal point of view, the difference between the two concepts is blurred and only a "right to explanation" has been theorized so far (Section 3). However, the new AI Act proposal, in its original version, provided some important transparency principles that might influence the way in which AI systems are built from a technical point of view (Section 4). In the last section of this article, the existence of a "right to technical interpretability" is theorized.

2. The concepts of "technical interpretability" and "explainability"

Before examining the different sources of law that contain the right of explanation, it is beneficial to consider about the distinction between explainability and interpretability. Although some researchers have attempted to provide a thorough

definition of the two words[1], there is still a lot of ambiguity in the literature regarding these notions, and they are frequently used indiscriminately[2]. The two terminologies are treated as separate notions in this essay. If we consider the definition provided by Rudin (2022)[3], explainability occurs when "a second (post hoc) model is created to explain the first black box model". In this scenario, there are two models: the first is opaque since it is impossible to understand why it produced a particular output, and the second, which produces a human-intelligible output, is used to attempt to piece together the factors that led to the first model's output. Explainability, then, is the attempt to explain the output of a black box.

According to the Author, in certain circumstances explainability may not be a reliable method, because explainable machine learning techniques offer explanations that are not accurate representations of what the original model computes, and therefore suggest an incorrect explanation. In fact, continues the author, they are unable to accurately reconstruct the original model. If the explanation accurately described what the original model computed, it would be identical to the original model, negating the very necessity for the original model in the first place: in other words, the second model would be interpretable[4].

Babic et al. (2021)[5] also present some doubts, pointing out three circumstances: first, as it is frequently the most accurate, the opaque function of the black box continues to serve as the foundation for AI/ML choices. Second, there cannot be a perfect approximation between the white box and the black box; otherwise, there would be no distinction between the two (as noted by Rudin). Additionally, fitting the black box, frequently just locally, is the main priority rather than accuracy. Finally, the justifications offered are post hoc. This draws attention to a crucial point that should be taken into account: based on

---

* Postdoctoral Researcher, Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands; Department of Mathematics and Geosciences, University of Trieste; School of Engineering, Carlo Cattaneo University – LIUC, Varese, Italy. Funded by the REMIDE project, Carlo Cattaneo University – LIUC and by the UNI 4 JUSTICE Project.

1 Köhl M.A., et al. (2019) 'Explainability as a non-functional requirement', *2019 IEEE 27th International Requirements Engineering Conference (RE)*, IEEE, pp. 363-368; Chazette, L., Schneider, K. (2020) 'Explainability as a non-functional requirement: challenges and recommendations', *Requirements Engineering,* vol. 25(4), pp. 493-514; Chazette, L., Brunotte, W., Speith, T. (2021) 'Exploring explainability: A definition, a model, and a knowledge catalogue', *IEEE 29th International Requirements Engineering Conference (RE)*, IEEE, pp. 197-208.

2 Rudin, C., et al. (2022) 'Interpretable machine learning: Fundamental principles and 10 grand challenges', *Statistics Surveys,* vol. 16, pp. 1-85.

3 Rudin, C. (2019) 'Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead', *Nature Machine Intelligence,* vol. 1(5), pp. 206-215.

4 Ibid.

5 Babic, B., *et al.* (2021) 'Beware explanations from AI in health care', *Science,* vol. 373(6552), pp. 284-286.

current knowledge, black box models are typically more accurate than alternative models. Nowadays, deep learning outperforms white box models in a number of fields, including imaging. But in some circumstances, interpretable models might be just as accurate as black boxes[6]. Petch et al.[7] suggest a general guideline for choosing whether to utilize black boxes:

> […] data scientists should train models using both interpretable and black-box methods to assess whether there is, in fact, an accuracy vs interpretability tradeoff in the specific case on which they are working. If there is no meaningful difference in accuracy between an interpretable model and a black box, an interpretable method should be used. However, if a black-box model does provide a higher degree of accuracy, the stakes of the decision should be considered. If the decision that will be informed by the model is a relatively low stake, a small improvement in accuracy may justify the use of a black box. However, if the stakes are high, it is reasonable to require a greater improvement in accuracy before sacrificing interpretability. Ideally, gains in accuracy from black-box methods should be sufficient to translate into meaningful improvements in clinical outcomes such as reduced morbidity or mortality. If the use of a black box model can be justified, explainability techniques should be employed to make the model and its predictions as transparent as possible, but clinicians should be aware of their limitations and be cautious of overinterpreting, which can lead to narrative fallacies.

The term interpretability describes a model that is understandable in a way that enables people to comprehend the substantial (not mere technical) justification that resulted in a specific output. "Justifiable" could be used as a synonym here; someone must be able to understand the logic involved in the decision, whether they be the modeler, the domain expert, the final user, or the individual who will be affected by the outcome. It is important to note that a model's mere mathematical justification and comprehension of the model structure (such as being able to look at the source code) are insufficient for humans to comprehend why the model came to a particular conclusion. Knowing the source code of a neural network, the number of parameters (weights), layers, or features, without knowing the reason why a relationship between data is found, may not enable humans to understand the output.

Consider a classification algorithm that categorizes upcoming patients of a particular condition into healthy or unhealthy based on a data set comprising millions of physiological characteristics of past patients. In order to determine whether the system classifies a new patient as healthy or ill, doctors may enter data from the patient, including blood levels, symptoms, anamnesis, genomic

6 Christodoulou, E., et al. (2019) 'A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models', *Journal of clinical epidemiology,* vol. 110, pp. 12-22.

7 Petch, J., Di, S., Nelson, W. (2021) 'Opening the black box: the promise and limitations of explainable machine learning in cardiology', *Canadian Journal of Cardiology.*

information, lifestyle choices, age, number of children, ethnicity, weight, height, number of sleep hours, job, place of birth, etc. However, because of the numerous and intricate features, parameters, and layers that are employed in producing the output, the system is unable to determine the cause of the patient's illness, such as the fact that the blood levels are abnormal for someone of the patient's age, ethnicity, weight, and daily amount of exercise. It's even possible that the system depended on unnecessary features by coincidence, such as the ID number or the date of admission to the hospital. Nor the physician nor the modeller has any mean to determine this. Das and Rad[8] consider that:

> the large number of parameters in Deep Neural Networks (DNNs) make them complex to understand and undeniably harder to interpret. Regardless of the cross-validation accuracy or other evaluation parameters which might indicate a good learning performance, deep learning (DL) models could inherently learn or fail to learn representations from the data which a human might consider important. […] Hence, often the ability to interpret AI decisions are deemed secondary in the race to achieve state-of-the-art results or crossing human-level accuracy.

In addition, Rudin[9] notes that interpretability is not a black and white concept, but rather a spectrum:

> there is a spectrum between fully transparent models (where we understand how all the variables are jointly related to each other) and models that are lightly constrained in model form (such as models that are forced to increase as one of the variables increases, or models that, all else being equal, prefer variables that domain experts have identified as important[...]).

There are differences in the literature about *what* should be explainable and in what context. Chazette et al.[10] summarize the elements that, according to the literature and their own analysis, should be explained, such as the inference processes for certain problems, the relationships between the inputs and outputs, parameters and data structures, intentions, behaviors in real-world, underlying criteria for the decision, predictive accuracy, and user preferences. Often, what is meant with the term "Exaplainable Artificial Intelligence (XAI)" is instead "Intelligible AI". The concept of "intelligibility" is crucial because it encompasses a wide range of considerations that need to be made, including cultural differences[11], mental

8 Das, A., Rad, P. (2020) 'Opportunities and challenges in explainable artificial intelligence (xai): A survey'. Available at: https://arxiv.org/abs/2006.11371.

9 Rudin, C. (2019) 'Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead', *Nature Machine Intelligence*, vol. 1, pp. 206-215.

10 Chazette, L., Brunotte, W., Speith, T. (2021) 'Exploring explainability: A definition, a model, and a knowledge catalogue', *IEEE International Requirements Engineering Conference.*

11 Including language: different languages may have different ways of expressing a concept. Localization is an important element of the transparency principle.

capacity, age, educational attainment, experience and expertise, preferences in visualization and design, and numerous other variables that may affect a recipient's capacity to comprehend a given output. As stated in Article 13 of the GDPR, it is essential to adapt to the addressee (the audience), and this is especially true when there is an impact on their life. Numerous scholars have emphasized the significance of the addressee's comprehension, which might vary depending on the circumstances[12]. However, in this article we will only focus on the inherent technical intelligibility of decisions, not on all the other elements that make an output intelligible, as important.

For instance, considering a system that predicts the likelihood of not being able to pay back a mortgage and is used by a financial institution to deny credit, we would consider it interpretable only if it made clear which financially significant factors – such as wage, job type, age, concurrent loans, marital status, and education – were used by the model to produce the output, what relationship were found between them (e.g., educated persons are more likely to have high incomes), and which ones were given a higher weight than others (for example, the system could weight the past mobility as an unfavourable condition and weight it more that an advanced age). Even if this explanation were to be given in mathematical terms, the modellers would still be able to translate it such that bank employees could understand it, and the personnel would then be able to explain it to the mortgage applicant in plain language.

3. The transparency principle and the "right to explanation"

Transparency is a key principle and an overarching obligation in the whole EU legislation and in particular within the Digital Strategy, but it is also an important ethical and legal requirement provided by national laws and guidelines in some fields relating to high-risk systems.

The "right of explanation" in GDPR is part of transparency: data subjects have the right to receive information about the rationale behind or the criteria

12 Ribeiro, M.T., Singh, S., Guestrin, C. (2016) "'Why should I trust you?" Explaining the predictions of any classifier', *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135-1144; Carvalho, D.V., Pereira, E.M., Cardoso, J.S. (2019) 'Machine learning interpretability: A survey on methods and metrics', *Electronics,* vol. 8(8), p. 832; Miller, T. (2019) 'Explanation in artificial intelligence: Insights from the social sciences', *Artificial intelligence,* vol. 267, pp. 1-38; Rosenfeld, A., Richardson, A. (2019) 'Explainability in human–agent systems', *Autonomous Agents and Multi-Agent Systems*, vol. 33(6), pp. 673-705; Barredo Arrieta, A., et al. (2020) 'Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI', *Information fusion*, vol. 58, pp. 82-115; Chazette, L., Brunotte, W., Speith, T. (2021) 'Exploring explainability: A definition, a model, and a knowledge catalogue', *IEEE International Requirements Engineering Conference.*

relied on in reaching an automated decision that has an impact on their life, and about the significance and envisaged consequences of the processing of their data, as provided by Articles 13 and 14 of GDPR. Controllers must provide meaningful information about the logic involved in the decision process, not necessarily a complex explanation of the algorithms used or the disclosure of the full source code[13], but a "sufficiently comprehensive explanation that allows the data subject to understand the reasons for the decision"[14]. This concept is closer to interpretability than to explainability: it is more important that data subjects can understand what a model did than how it did it from a technical point of view. The right of explanation is also present in the Council of Europe's Convention 108+, with a broader application than in GDPR, as explained in the Explanatory Report, in Article 10:

> data subjects should be entitled to know the reasoning underlying the processing of their data, including the consequences of such reasoning, which led to any resulting conclusions, in particular in cases involving the use of algorithms for automated decision-making including profiling. For instance, in the case of credit scoring, they should be entitled to know the logic underpinning the processing of their data and resulting in a "yes" or "no" decision, and not simply information on the decision itself. Without an understanding of these elements, there could be no effective exercise of other essential safeguards such as the right to object and the right to complain to a competent authority.

Transparency means that Controllers must provide to data subjects (e.g., patients) "relevant information related to fair processing, communicate and facilitate the exercise of their rights, enabling them to understand, and if necessary, challenge the data processing"[15]. However, the current legislation does not dictate the provision to data subject of a "technically faithful explanation", since not all models allow the modeller or the user to know the true reasoning behind the output. A legally compliant explanation could also be that a job applicant was rejected "because the CV does not match the minimum requirements listed in the job posting in terms of experience and education", while the truth is that the true reason is not known, or that it is only partially known but it is full of hidden biases[16]. Since the reasons are not correctly communicated to the job applicants, they would have no knowledge of the biases and could not file a discrimination lawsuit.

---

13 Even consumer law does not require full disclosure of all the algorithms involved. Directive (EU) 2019/2161 requires transparency only regarding the main parameters used by the model.

14 Article 29 Data protection Working Party. Guidelines on Automated individual decisionmaking and Profiling for the purposes of Regulation 2016/679. WP215 1 (2017).

15 Article 29 Data protection Working Party. Guidelines on Transparency under Regulation 2016/679. WP260 rev. 1 (2018).

16 For a broader analysis on hiring software discrimination, see Kelly-Lyth, A. (2021) 'Challenging biased hiring algorithms', *Oxford Journal of Legal Studies*, vol. 41(4), pp. 899-928.

It must be noted that GDPR and national law also provides that the data subjects must express their informed consent. How is possible to truly and freely express consent if it is not known why the system has given a certain output? As we have seen above, only interpretability can guarantee that the substantial reasons on which the decision has been relied are known.

4. Articles 13 and 14 of the AI Act Proposal

In April 2021, the European Commission published the AI Act proposal, since the specific characteristics of certain AI systems may have an impact on user safety and fundamental rights, creating new risks which need to be addressed. The most important innovation of the proposal is the establishment of four risks categories for AI systems, in order to protect citizens' fundamental rights and freedoms[17]. The risk categories are related to the degree (intensity and scope) of risk for the safety or fundamental rights of citizens. Taking inspiration from the product safety legislation, the classification of risks is based on the intended purpose and modalities for which the AI system is used, not only on their specific function. According to the new legal framework, some AI systems are considered as "high-risk" – in particular, AI decision support systems having an impact on important personal interests, e.g., in the case of healthcare – and must, therefore, fulfil new requirements before being put into the market or into service, including a risk management system, appropriate data governance measures and a quality management system, the use of high-quality datasets, the establishment of relevant documentation to enhance traceability, the sharing of adequate information with the user, the design and implementation of appropriate human oversight measures, and the achievement of the highest standards in terms of robustness, safety, cybersecurity, and accuracy, as well as the respect of applicable laws and regulations (e.g., GDPR). Article 13[18] use the phrase "enable users to interpret the system's output" and mention the concept of transparency twice. We need to understand what the legislator intended to

17 The explanatory memorandum attached to the proposal, in fact, notes that "The use of AI with its specific characteristics (e.g. opacity, complexity, dependency on data, autonomous behaviour) can adversely affect a number of fundamental rights enshrined in the EU Charter of Fundamental Rights […] In case infringements of fundamental rights still happen, effective redress for affected persons will be made possible by ensuring transparency and traceability of the AI systems coupled with strong ex post controls".

18 "Transparency and provision of information to users 1. High-risk AI systems shall be designed and developed in such a way to ensure that their operation is sufficiently transparent to enable users to interpret the system's output and use it appropriately. An appropriate type and degree of transparency shall be ensured […] 3. The information referred to in paragraph 2 shall specify: […] (d) the human oversight measures referred to in Article 14, including the technical measures put in place to facilitate the interpretation of the outputs of AI systems by the users […]".

achieve through this wording. Analysing the heading of article 13, we find the words "Transparency" and "information", liked by the conjunction "and", a sign that the two concepts have separate meanings. Therefore, the law considers transparency as being a different concept than merely providing information to users, such as in the readme text file or in the technical instructions to users. However, this is not enough to say that a right to technical interpretability is enshrined in the AI Act. In order to analyse the meaning of Article 13, we also need to have a look at the related recital. However, Recital 47 is not very helpful in providing information about Article 13, since it only adds a few specifications to its text:

> To address the opacity that may make certain AI systems incomprehensible to or too complex for natural persons, a certain degree of transparency should be required for high-risk AI systems. Users should be able to interpret the system output and use it appropriately. High-risk AI systems should therefore be accompanied by relevant documentation and instructions of use and include concise and clear information, including in relation to possible risks to fundamental rights and discrimination, where appropriate.

In addition, it seems to link the concept of interpretability to the mere provision of documents and instruction.

Article 14 mentions the concept of interpretability when referring to the human oversight measures, prescribing that one of the measures to achieve it is to enable the user to "correctly interpret the high-risk AI system's output, taking into account in particular the characteristics of the system and the interpretation tools and methods available". Interpretability is then a mandatory, yet alternative, measure to make sure that a human is always kept in the loop to oversee the behavior of the AI system. Although this provision is not, alone, sufficient to affirm that a right to technical interpretability exists in the AI Act, it is certainly a strong argument in its favor. Other recitals may contain relevant information regarding the degree of transparency required for high-risk systems: for example, recital 39 states that:

> […] The accuracy, non-discriminatory nature and transparency of the AI systems used in those contexts are therefore particularly important to guarantee the respect of the fundamental rights of the affected persons, notably their rights to free movement, non-discrimination, protection of private life and personal data, international protection and good administration.

One might argue that without technical interpretability, or at least a very accurate explainability technique, it is impossible to guarantee the absence of discriminatory outputs and provide remedies against them.

The AI Act, overall, does not respond to our question regarding the existence of a "right to technical interpretability". The matter should be looked under the lens of a systematic interpretation, taking into consideration the sources of law that protect fundamental human rights.

5. The "Right to Technical Interpretability" as a fundamental right

Many academics have argued that black box algorithms shouldn't be used as normal practice in industries like medicine because of their internal opacity. This is because they can't ensure key aspects of good medical treatment[19]. Das et al.[20] believe that, due to the significant impact of data bias, trustworthiness, and adversarial examples in machine learning, it is currently not recommended to blindly accept the output of a highly predictive classifier.

Regarding post-hoc explainability, although it may be a useful tool, the limitations must be taken into account, as explained by Rudin[21] and Babic[22]. Vale et al. illustrate the limitations of post-hoc explainability techniques in proving discrimination, arguing that the tendency towards showing result parity that is necessary for EU non-discrimination law is absent from post-hoc explainability methodologies[23]. They argue that, because of their technical flaws, they are occasionally unstable and exhibit low fidelity, being unable to convincingly prove that there is no discrimination: the limited bias types discovered by post-hoc explainability approaches require their use to be contextualized and limited. According to them, the use of post-hoc explainability approaches is beneficial, particularly in the creation and development of models, but they might not be appropriate for use in regulatory or legal applications; as a result, they cannot be promoted as panaceas and cannot be valued solely in a vacuum without regard to more comprehensive fairness metrics. The authors believe that the substantive legal weight that post-hoc explainability procedures might be able to provide is questionable if they cannot establish *prima facie* discrimination. Therefore, they reach the same conclusion as Rudin and Babic: if a black-box model's

19 Rudin, C. (2018) 'Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead', *Nature Machine Intelligence*, vol.1; Kundu, S. (2021) 'AI in medicine must be explainable', *Nature Medicine*, vol. 27(8), pp. 1328-1328.

20 Das, A., Rad, P. (2020) 'Opportunities and challenges in explainable artificial intelligence (xai): a survey'. Available at: https://arxiv.org/abs/2006.11371.

21 Rudin, C. (2018) 'Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead', *Nature Machine Intelligence*, vol. 1.

22 Babic, B., et al. (2021) 'Beware explanations from AI in health care', *Science* 373.6552, pp. 284-286.

23 Vale, D., El-Sharif, A., Ali, M. (2022) 'Explainable artificial intelligence (XAI) post-hoc explainability methods: Risks and limitations in non-discrimination law', *AI and Ethics*, pp. 1-12.

insights and/or internal workings cannot be relied upon, it should not be used in situations where its judgments could have significant and/or long-lasting implications. Another factor that encourages the use of interpretable models in high-risk systems, and more broadly in applications that may have a significant impact on citizens' life, is the fact that these systems frequently affect fundamental human rights, which are safeguarded by both international legal instruments and most Constitutions. This suggests that even a small chance of discrimination resulting from unintentional bias is not accepted by the legal system. Black boxes make it impossible to regulate the model output and to examine the reasoning process to see whether it is based on unfair or irrelevant criteria. Additionally, it would be exceedingly challenging to determine whether the prejudiced output reflects societal biases or the modeler's own unconscious biases or opinions, which could cause issues with determining responsibility in some legal systems (e.g., for gross negligence or wilful misconduct in creating a biased model)[24]. Although neither GDPR nor the AI Act unambiguously dictates the use of interpretable techniques, and some authors have even been challenging the very existence of a "right to explanation"[25], in many AI applications the only way to protect citizen's fundamental rights and freedom is to employ interpretable models. Taking into account the systematic interpretation of the EU and international legal framework surrounding high-risk systems, it is possible to argue that interpretability should be used as a standard in those fields (and even in other sensitive fields) and that black boxes should only be used in situations where it is possible to make a decision by evaluating factors other than the AI output. The very possibility of expressing informed consent and challenging the decision made on the basis of an automated decision-making system is excluded by the opacity and complexity of black boxes. The lack of technical interpretability prevents the exercise of many fundamental rights, such as the right to a fair trial, to self-determination, to non-discrimination, and more. A "right to technical interpretability" should, therefore, be theorized at the European level, being considered a fundamental right, and embedded in the AI Act proposal.

24 It is worth noting that the European Commission has recently published a proposal for a directive on AI Liability, which covers the civil liability connected to high-risk systems. The topic is, however, to broad to be addressed here. For more insights on AI liability, see Gallese, C. (2022) 'Suggestions for a Revision of the European Smart Robot Liability Regime', *Proceedings of the 4th European Conference on the Impact of Artificial Intelligence and Robotics (ECIAIR)*, vol. 4(1), pp. 29-35.

25 Wachter, S., Mittelstadt, B., Floridi, L. (2017) 'Why a right to explanation of automated decision-making does not exist in the general data protection regulation', *International Data Privacy Law*, vol. 7(2), pp. 76-99.

Chapter XX
Innovative Versus Recurrent Perspectives on the Liability for Autonomous and Incorporated Artificial Intelligence

by Juanita Goicovici*

INDEX: 1. Establishing the contours of the AI manufacturers/importers and distributors liability. – 2. Recurrent premises and epitomical features of liability for defective products. Epitome of the "risks/benefits" test. – 3. Tripoded taxonomy of AI design defectiveness – AI manufacturing flaws – AI informational cadences. – 4. Innovative approaches: targeted rebuttable presumption of causality in the case of self-evolving AI. – 5. Proportionality exigencies for issuing procedural orders on inculpatory evidence disclosure. – 6. Errors in the identification of the AI producer: an aporia that retaliates the delimitation of the sphere of responsible persons. – 7. Subrogation in the injured person's right to compensation and adjustment of personal sphere of incidence. – 8. Conclusive remarks.

1. Establishing the contours of the AI manufacturers/importers and distributors liability

Autonomous and incorporated AI engaged in prejudicial episodes continuously raised interrogations for legal practitioners, in terms of establishing the contours of the manufacturers/importers and distributors liability, while contrasting the versions of subjective liability, in the perimeter of AI-incorporating products, when exploring clichés and recurring features resulting from the juxtaposition over pre-existing liability regimes[1], which are currently absorb-

* Associate lecturer at the Private Law Department of the Babeş-Bolyai University of Cluj-Napoca, Romania. Her predominant areas of research are encompassing Digital Contracts, Consumer Protection Law, Liability for defective/endangering AI, Civil Liability for the Processing of Personal Data, B2C Contracts and Specific Asymmetries. She holds a PhD degree in the progressive formation of B2B and B2C contracts and an advanced studies degree in Business Contracts Law, at the Babeş-Bolyai University of Cluj-Napoca. She authored numerous articles on Consumer Law and on Business Contracts, addressing the ownership of Movables in the Category of Autonomous and Embedded Artificial Intelligence Civil Liability for Damages Caused by AI Defectiveness.

1 Yap, J. Q., Lim, E. (2022) 'A Legal Framework for Artificial Intelligence Fairness Reporting, NUS Law Working Paper No. 007', *Cambridge Law Journal*, vol. 81. Available at: https://ssrn.com/abstract=4128641; Zhang, Y., Gosline, R. (2022) 'Understanding Algorithm Aversion:

ing the attention of the EU's deciding legislative bodies[2]. As noticeable[3] in the intricated tapestry of the three regulatory proposals intensely debated, namely the Proposal for Directive on AI liability COM/2022/496 (Drafted AILD), the Proposal for the Artificial intelligence Act COM/2021/206, respectively the Proposal for a Directive on liability for defective products COM/ 2022/495, which would supersede the rules in the current regulation of producer/importer liability for defective products Directive (EC) 85/374, the latter being the most longevous of the characteristic regulations for the pillars of European consumer law – the range of appropriate legal instruments, which are specific to civil liability for damages caused by interacting with defective products involving autonomous or embedded AI are diversifying, becoming crucial benchmarks for solving issues such as those relating to accessing pertinent evidence for litigiously disputed AI defectiveness.

In this sense, as we will have the opportunity to observe, the Drafted AILD[4] firstly reduces for the consumer/claimant the difficulties related to the burden of proof, by using mechanisms such as the obligation to disclose information[5] imposed on the manufacturer/importer/supplier of the AI products hypostasised in the procedural position of the defendant; secondly, the rebuttable legal presumptions of causation is establishing, in favour of consumers seeking compensation for damages caused by defective AI[6], the possibility of obtaining information from the manufacturer/designer regarding the manifested flaws

When Do People Abandon AI After Seeing It Err?', *MIT Sloan Research Paper*, no. 6846-22. Available at: https://ssrn.com/abstract=4299576, http://dx.doi.org/10.2139/ssrn.4299576.

2 De Bruyne, J., Dheu, O., Ducuing, C. (2022) 'The European Commission's Approach to Extra-Contractual Liability and AI – an Evaluation of the AI Liability Directive and the Revised Product Liability Directive'. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4239792; Burak, M.F. (2022) *Effects of Artificial Intelligence on E-Commerce*, in A. N. Özker (edt.)(2022) *Reviews in Administrative and Economic Science Methodology, Research and Application,* Ed. Livre de Lyon Publishing, Lyon, pp. 91-100. Available at: https://ssrn.com/abstract=4315366.

3 Waltermann, A. (2021) 'On the legal responsibility of artificially intelligent agents. Addressing three misconceptions', *Technology and Regulation*, pp. 35-43. Available at: https://techreg.org/article/view/10985/11959.

4 Goicovici, J. (2023) 'Rebuttable Presumptions of Causality and Reverberations of Evidence Disclosure, as Epitomic Pieces in the Physiognomy of Liability for Defective AI', *Journal of Law, Market & Innovation*, vol. 28, pp. 28-58. Available at: https://ojs.unito.it/index.php/JLMI/article/view/8892.

5 Rodríguez de las Heras Ballell, T. (2023) 'The revision of the product liability directive: a key piece in the artificial intelligence liability puzzle', *ERA Forum*, vol. 24, vol. 2, pp. 247-259. Available at: https://doi.org/10.1007/s12027-023-00751-y.

6 Raposo, V.L. (2022) *The European Draft Regulation on Artificial Intelligence: Houston, We Have a Problem*, in Marreiros, G., Martins, B., Paiva, A., Ribeiro, B., Sardinha, A. (eds.), *Progress in Artificial Intelligence*. EPIA, Springer, Cham, pp. 66-73. Available at: https://doi.org/10.1007/978-3-031-16474-3_6.

of AI systems with a high degree of risk, which were registered/documented following the prerequisites set under specific legislation[7].

The triptych of categories of defects which are covered by the material scope of application of the special liability of AI manufacturers/importers is compartmented as follows:

- faults occurring in the manufacturing chain (human error/artificial intelligence, algorithms, industrial equipment failures, etc.;
- AI design defects, exaggerated by the manner under which it has been exploited by the user (the "risk-utility balance");
- the extra-contractual liability connected to untransparent use of automated decision-making systems (ADM explainability and transparency).

We argue that the lacunary or erroneous decisions / biased outputs, as well as low-resiliency against attempts to alter autonomous / product-incorporated AI use or performance may be preferably conceived as product dysfunctionalities, while permitting the consumer to engage the producer's or retailer's liability[8] for specific damage (the dichotomous concepts of "autonomous material damage vs. derivative material damage" or the dichotomic categories of "liability for product security deficiencies" vs. artificial intelligence "design flaws")[9].

7 Veale, M., Borgesius, F.Z. (2021) 'Demystifying the draft EU artificial intelligence act – analysing the good, the bad, and the unclear elements of the proposed approach', *Computer Law Review International*, vol. 22, vol. 4, pp. 97-112. Available at: https://www.degruyter.com/document/doi/10.9785/cri-2021-220402/html.

8 Butler, A. (2018) 'Products Liability, and the Internet of (Insecure) Things: Should Manufacturers Be Liable for Damage Caused by Hacked Devices?', *University of Michigan Journal of Law Reform*, vol. 50(4). Available at: https://ssrn.com/abstract=2955317.

9 Cabral, T. S. (2020) 'Liability and artificial intelligence in the EU: Assessing the adequacy of the current Product Liability Directive', *Maastricht Journal of European and Comparative Law*, vol. 27(5), pp. 615-635; Cabral, T. S. (2018) 'Robotics and AI in the European Union: opportunities and challenges', *UNIO - EU Law Journ*al, Centro de Estudos em Direito da União Europeia (CEDU), vol. 4(2), pp. 135-146; Cabral, T. S., Kindylidi, I. (2021) 'Sustainability of AI: the case of provision of information to consumers', *Sustainability Journal*, vol. 13(21); Calo, R. (2017) *Artificial Intelligence Policy: A Primer and Roadmap*. Available at: https://ssrn.com/abstract=3015350; Chatzipanagiotis, M. P. (2020) 'Product Liability Directive and Software Updates of Automated Vehicles', *Proceedings of SETN 2020 – 11th Hellenic Conference on Artificial Intelligence*. Available at: https://ssrn.com/abstract=3759910; Chatzipanagiotis, M. P., Leloudas, G. (2020) 'Automated Vehicles and Third-Party Liability: A European Perspective', *University of Illinois Journal of Law, Technology & Policy*. Available at: https://ssrn.com/abstract=3519381.

2. Recurrent premises and epitomical features of liability for defective products. Epitome of the "risks/benefits" criterion

When assessing the AI defectiveness[10], the dynamics of the "classical" benchmarks set for torts liability might prove to be insufficient to meet the hypercomplex legal intricacies associated to the insidious, progressively generalized presence of autonomous/embedded AI. These gradual "metamorphoses" of the liability regimen are characterized by intriguing traits in ameliorating the burden of proof and challenging the legitimacy of AI producers' tactics of recurring to untransparent or even contradictory information on the non-conformity of embedded AI or AI components. National courts might find that there is a non-compliance with an obligation of due diligence on the manufacturer's side or on the one who has used the AI system, when, assessing the circumstantial evidence, it would appear plausible that the culpable omission of the AI user considerably influenced the compromised result generated by the AI system. On another versant of the discussion, in judicial practice, salient interrogations were proclaimed addressing the pertinence of the criteria extracted from "classicized" torts law, under national regulations, for establishing the contours of the AI user's culpable conduct as a legal basis for retaining the latter's faulty-based liability, in the context where the current provisions of EU regulations are sibylline in this respect; these aspects are curatorial for an unfashionable aporia, pertaining to the emblematic issues of specific liability of manufacturers or users for AI deficiencies the manifestation of which has not been triggered. Saliently, the "serious risk" criterion is recently used to describe a risk for which, based on a risk assessment and considering the normal and foreseeable use of the product, the combination of the probability of the occurrence of a danger of injury and the degree of severity of said injury requires a judicial (reparatory) intervention, albeit not immediately obvious.

At the "nadiral point" in the horizon of the civil liability for deficient AI response, as outlined in article 7 of the Drafted AILD, it is noticeable that it addresses the potential faulty-based liability cases involving the malevolent use or interaction of the user (in the sense of natural person or legal person using the disputed product in a controllable manner, either predominantly extra-professional or in the context of providing professional services), interfering with the actions/omissions of algorithmic systems or deficient autonomous/incorporated AI[11], However, when partially "imitating" several traits of the torts

10 Cabral, T. S. (2018) 'Robotics and AI in the European Union: opportunities and challenges', *UNIO - EU Law Journa*l, Centro de Estudos em Direito da União Europeia (CEDU), vol. 4(2), pp. 138-139.

11 Dastani, M., Yazdanpanah, V. (2023) 'Responsibility of AI Systems', *AI & Society*, vol. 38(2), pp. 843-852. Available at: https://doi.org/10.1007/s00146-022-01481-4; Davida,

liability regime, in the specific liability the regulation of which is detailed in the Drafted AILD there is a ostensible persistency of the subjective criterion relating to the faulty choices or inculpatory behavior of the AI users. In fact, the "adapted" liability regimen do not "dramatically" distort the pre-existing regimes under the national regulations pre-existing in the legislative framework of the member states, at the level of mandatory or suppletory norms addressing the coverage of damages caused by the illegal/imputable or inexcusable action/omission of the users. Nevertheless, the "adapted" structure of the civil liability for AI "distorted result" is conceived as an embodiment of the "classicized" matrix of torts liability and "dilutes" its reverberations on the role attributable to the culpable behavioral traits that were imputable to the AI user.

3. Tripoded taxonomy of "AI design defectiveness – AI manufacturing flaws – AI informational cadences"

Recurrently echoed in the conceptual areal describing the liability for defective AI[12], the reference to the faulty-based assessment of the AI user's/supplier's conduct might contribute to the altering of the effectiveness of non-contractual civil liability remedies available to injured consumers[13], instead of fortifying it, particularly when intertwining the mentioned subjective criterion of faulty behavior with the rebuttable presumption of good-faith and the rebuttable presumption of innocent (unimputable) conduct characterizing the "classicized" regimen of torts liability. For the pertinency of the subjective criterion of imputable misconduct to be noticeable, an adequate understanding of the "concentric circles" involved in the structure of the extra-contractual liability for damages caused to consumers by AI products is imperatively demanded, when "dissipating" the latent possibility of resorting to the invoking the "development risk", as a causal exoneration from civil liability, since the text of article 6, 1st para., let. (c) of the Draft Directive COM/2022/495, 2022/0302 (COD)

Z. (2021) 'Chatbots by business vis-à-vis consumers: A new form of power and information asymmetry', *SHS Web of Conferences*, vol. 129, article 05002129. Available at: https://doi.org/10.1051/shsconf/202112905002 and https://www.shs-conferences.org/articles/shsconf/abs/2021/40/shsconf_glob2021_05002/shsconf_glob2021_05002.html.

12 Corbin, B. (2019) 'When Things Go Wrong: Redefining Liability for the Internet of Medical Things', *South Carolina Law Review*, vol. 71(1). Available at: https://ssrn.com/abstract=3375070; Diamantis, M. (2022) *Vicarious Liability for AI*, in Johnson K., Reyes C. (eds.) (2022) *Cambridge Handbook of AI and Law*. Cambridge, University of Iowa Legal Studies Research Paper No. 2021-27. Available at: https://ssrn.com/abstract=3850418; Diamantis, M. (2020) 'Who Pays for AI Injury?', *Oxford Business Law Blog*. Available at: https://ssrn.com/abstract=3592546.

13 Ebers, M. (2021) 'Liability for Artificial Intelligence and EU Consumer', *Journal of Intellectual Property, Information Technology and Electronic Commerce Law*. Available at: https://ssrn.com/abstract=3855110.

expressly addresses the cases of product defectiveness for which the liability attributable to the producer/importer is envisaged, particularly in situations in which the prejudicial product is regimented in the "self-learning" category. The mentioned provisions tend to specify that the intrinsic defectiveness of a product might be assessed particularly by considering the dynamics of continuously evolving circumstances, such as the product ability to continue learning after its practical implementation. At the antipode of the reverberations that could be retained within the areal of cases where the AI user reasonably/vigilantly conducted examination of the instructions accompanying the AI product prior to the engaging in the interaction, manufacturers and users may be held liable within the substantial sphere of the extra-contractual liability for the distorted responses of AI systems regimented in the self-learning or self-training category, which are self-trained from interaction with users consumers' in the latter's cases, the debate is centered on the necessity of tailoring the users' responsibility for the prejudicial AI systems the constant adaptation of which constitutes the result of "self-learning abilities" from its interaction with the AI users, covered by this type of liability including "self-acquired" AI defects and deficiencies manifested throughout the self-evolving AI cycle.

4. Innovative approaches: targeted rebuttable presumption of causality in the case of self-evolving AI

In the hypotheses[14] when the harm occurred due to the circumstantial change of the conditional elements that were incidental in the perimeter of the AI responses[15], where lacking a genuinely adapted response to the variation of the conditional premises[16], legal interrogations may be raised on whether the AI producer/designer should have been considering the adjusting/adaptability of autonomous or incorporated AI to changing/evolving environmental parameters and to whether such an omission may represent an autonomous design defectiveness. This effect is even more accentuated for those AI defects

14 Sheriff, K. (2020) 'Defining Robotic Autonomy in the Context of Tort Liability', *Emory Law*. Available at: https://ssrn.com/abstract=3743478, http://dx.doi.org/10.2139/ssrn.3743478.

15 Geistfeld, M. (2021) 'Strict Products Liability 2.0: The Triumph of Judicial Reasoning over Mainstream Tort Theory', *Journal of Tort Law*, vol. 14(2), NYU School of Law, Public Law Research Paper No. 22-01. Available at: https://ssrn.com/abstract=3956019; Geistfeld, M. (2020) *Principles of Product Liability, Third Edition: 'Introduction' and 'Strict Products Liability 2.0'*, in Geistfeld, M. A. (2020) *Principles of Products Liability*, Foundation Press, 3rd ed., New York University Law and Economics Research Paper No. 20-36. Available at: https://ssrn.com/abstract=3691064.

16 Howells, G., Twigg-Flesner, C. (2022) *Interconnectivity and Liability: AI and the Internet of Things*, in di Matteo L., Poncibo C., Cannarsa M., Siren P. (eds.) (2022) *Artificial Intelligence: Global Perspectives on Law & Ethics*, Cambridge University Press. Available at: https://ssrn.com/abstract=3843134.

materialized during the interaction with end users, while not being applicable the requirement of the previous defect[17], which would have been characteristic of the "classical" warranty regime incumbent on the seller for the hidden defects of the product[18]. Saliently, it is the "constant monitoring" obligation incumbent on AI producers/suppliers that represents the piece of resistance in the discussion centered on establishing the conditions of civil liability for bodily/patrimonial damages affecting consumers as result of dysfunctional response of AI systems regimented in the self-learning category.

5. Proportionality exigencies for issuing procedural orders on inculpatory evidence disclosure

Similarly to the case of the applying of the rebuttable presumption relating to the causal link between the defendant's imputable behavior or inexcusable negligence in reckless interacting with a controllable AI system, for the second species of rebuttable presumptions regarding the establishing of a "failure" (connected to faulty conduct) of the defendant's compliance to the obligation to engage in precautionary measures or specific precautionary protocols, national courts may resort to the issuing of an evidence disclosure order, should these procedural measures be assessable as "proportionate". As decipherable in the text of article 3, 4th para of the Drafted AILD, national courts are expected to limit the ordering of disclosure of evidence to covering hypotheses when the plaintiff's reasonable efforts were unavailing, fruitless, or impuissant in accessing relevant evidence on AI defectiveness. Finally, for the procedural condition regarding the causal link or the causality relationship between the damage recorded in the consumer's patrimony and the existence of the *lato sensu* AI defectiveness, as a general requirement, it is noticeable that, most of the time, it imposes a difficult task on the consumer, despite the fact that the legislator partially simplifies the burden of proof in terms of proving the perviousness of the manifested deficiencies; proving the causal link remains difficult, since the aforementioned provisions do not exempt the claimant from the requirement to prove that the damage was caused by the AI structural or intrinsic deficiencies or, congruently, that the damage was caused by the interaction with the defective product. Most of the national legislators were not establishing a presumption of causality that is operable within the scope of special liability for product security deficiencies; therefore, in situations of this

17 Schütte, B., Majewski, L., Havu, K. (2021) 'Damages Liability for Harm Caused by Artificial Intelligence – EU Law in Flux', *Helsinki Legal Studies Research Paper No. 69*. Available at: https://ssrn.com/abstract=3897839.

18 Hacker, P. (2022) 'The European AI Liability Directives – Critique of a Half-Hearted Approach and Lessons for the Future', *Computer Law & Security Review*, vol. 51. Available at: https://ssrn.com/abstract=4279796.

kind, from the perspective of alleviating the burden of proof, in the matter of establishing the causal link, one returns to the perspective of using rebuttable judicial presumptions.

6. Errors in the identification of the AI producer: an aporia that retaliates the delimitation of the sphere of responsible persons

Describable as a convoluted task, the identification of the categories of responsible entities for covering the prejudicial effects of AI conformity lapses oscillated between selecting the "actors" involved in the conceiving, implementing, manufacturing, and embedding the AI elements. For instance, when assessing AI lack of safety, inadequacy for consume or defectiveness, it may become hardly feasible or perniciously inadequate to develop synchronized standards for eclectic types of algorithmic entities as these are typically adjusted to the various taxonomies of AI products. Establishing the "genealogy" of a complex AI system represents a procedural task which might require the detangling of the roles played and postures impersonated by the multiple-party entities involved in product design, or in the pre-launching assessment of safety standards, thus becoming a provoking task for the prejudiced claimants in identifying the liable defendants. The second element of the definitory taxonomy of liable persons for deficient AI inputs/outputs brings together a suite of "tectonic plates" that most commonly overlap in practice, given that certain categories of AI users are liable on a faulty-based regime for altering the AI responses which prejudiced third-parties' interests. The extra-contractual liability of AI manufacturers, importers or, in the alternative, AI distributors for safety deficiencies resulting from *lato sensu* conceptual defects is juxtaposed to the already existing liability regimes, regardless of whether the latter are of a contractual or extra-contractual nature, the consumer still being able to opt for placing the complaint on the "classical" grounds of civil liability for faulty conduct or to engage the professionals' contractual liability, including to avail himself/herself of the warranty for the hidden defects of the sold product, to the extent that (in rare cases) there would be identified the existence of a contractual relationship permitting the prejudiced consumer to invoke the contractual nexus as a source of rights and obligations; in addition to this conclusion derived from the provisions of Drafted AILD, it should be noted that the latter did not aim from the start to remove the functionality of these "classic" liability regimes available under national regulations; yet, it aimed at outlining a special regime of responsibility that would function in parallel, "doubling" the remedial variants already established in the national law systems for engaging the liability of professionals in relations with consumers prejudiced by the interaction with unsafe

AI. Ostensibly, the need to regulate a specific regime in the matter of biological damage (including the version of moral and material damages accruing to this type of prejudice, resulting from damage to the bodily integrity, health or life of the consumer) was generated by the requirement to (at least, partially) protect the consumer against unpredictable AI behavior, considered to be crucial in importance in a general hierarchy of harms whose coverage, in the national legal systems, might be precarious when facing the complex legal challenges brought by the generalized "infiltration" of the AI presence.

7. Subrogation in the injured person's right to compensation and adjustment of personal sphere of incidence

Recognizable as "claimant" would not only be the consumer solicitating compensation for damages, when a damageable output of an AI system has been produced, but also a claimant who has been subrogated in the compensatory rights of a person who has been injured under the terms of the law or a contract, or who acts as procedural agent of the prejudiced consumers, valorizing a joint-representation mandate, as described in article 2, pt. (6) of the AI Act Proposal. Nonetheless, should the deficient nature of the product-embedded AI be intertwined with elements related to a third-party's fault, omission or culpable behavior, and the manufacturer had been held responsible towards the victim of the damage for its financial coverage, subsequently the defendant manufacturer/importer would be able to justify the admission of a regress action towards a previous detainer of the compensation rights. Passive solidarity is characteristic for this type of civil liability, being a joint liability from which it follows that, in the event that a plurality of legal relationships is retained as covered by the prejudiced consumer's action for compensatory remedies, when multiple persons are liable for the physical or patrimonial damage caused to the consumer, their liability might be jointly retained' the multiple-party liability is seen through the lens of the provisions of domestic law regarding the recourse actions available to the debtors of the claims in damages against other participants upstream of the production/distribution chain who were find jointly and severally liable for compensating the losses invoked by the complaining consumer.

8. Conclusive remarks

While the liability of AI manufacturers, importers or, alternatively, distributors for safety deficiencies resulting from *lato sensu* manufacturing defects is an objective liability, retainable independently of the faulty elements of conduct, AI users' liability remains an extra-contractual faulty-based type of legal

responsibility for prejudicial consequences of interacting with embedded/autonomous AI. Comparatively, while in the perimeter of the "classical" torts liability for prejudicial defects, a three-dimensional classification of the categories of covered damages could be identified, the new drafts of normative acts in the matter of liability for AI deficiencies adds a fourth category of damages, related to the loss or alteration of data, as stemming from the deficient action/omission of AI systems. The structural elements or the components of the taxonomy of damages included in the material sphere of the specific liability for the deficiencies manifested in autonomous/embedded AI categories of products can be detected on four levels: i) on the first level, it is the biological, bodily or physical damage caused to the consumer that can be compensated following claimant's request; ii) the second tier is reserved to the allocation of moral damages incurred to consumers; iii) on the tertiary level, the substantial sphere of incidence of the extra-contractual liability regime targets the coverage of patrimonial damage resulting either from the autonomous material damages, or from the reverberations attached to the physical damage (derivative material damages); iv) forth, the material scope of the fault-centric type of liability includes the damage (property/moral damages) resulting from the loss or alteration of personal/non-personal data (vehiculated on non-professional purposes by the consumer). As it appears as preeminent in pursuing the objective of establishing "adapted" liability rules, national courts may resort to ordering the defendants to provide evidence enabling the claimant to plead for compensation under a non-contractual civil tort action; the latter procedural effect is corelated to norms whose material/substantial scope of application would be predominantly (but not exclusively) represented by the litigious situations involving the defectiveness of the AI systems, within the perimeter of extra-contractual, faulty-based liability and compensatory actions.

Chapter XXI
Digitalisation of Justice in the EU, Challenges and Future Prospects

by Anastasia Nefeli Vidaki*

INDEX: 1. Introduction. – 2. Online dispute resolution. – 3. Electronic access to case-law and legislation. – 4. Electronic filing and service. – 5. Videoconferencing. – 6. Predictive justice. – 7. Profiling. – 8. Conclusion.

1. Introduction

The European Union has expressed in the Tallin Declaration its political commitment towards ensuring efficient, user-centric digital public services for citizens and businesses. Among them is the provision of unhindered access to justice and cross border judicial cooperation, which have been undergone digitalization. In the fields of civil, commercial and criminal matters, European Union (EU) has already taken some major legislative steps towards then transfer of judicial proceedings and administration of justice from the analogical to the digital world, accompanied by some practical initiatives. The recast of service of documents and taking of evidence Regulations, the e-CODEX Regulation, the Proposals on AI Act and digitalisation of cross-border judicial cooperation, the e-justice portal, are considered the most vital aspects of the EU digital agenda in judicial matters[1]. In the near future, online dispute resolution with the assistance of algorithms based on machine learning is expected to give birth to a hybrid process, in which the role of new technologies would be crucial.

* PhD Candidate/ Researcher at the Cyber and Data Security Lab (CDSL), which is part of the Research Group on Law, Science, Technology & Society (LSTS) at the Faculty of Law and Criminology of Vrije Universiteit Brussel (VUB). She completed her bachelor studies in Law in the Aristotle University of Thessaloniki. During her studies, she spent a semester at the Faculty of Law of University of Cologne and carried through an internship at the Permanent Representation of Greece to the EU. She has also obtained a MsC on Law and ICT at the University of Piraeus (with distinction) and a LLM on Sociology of Law, Science and Technology at the National and Kapodistrian University of Athens (with distinction). She is a qualified lawyer and member of the Bar Association of Piraeus, Greece. She has fulfilled the Bluebook traineeship at the Justice and Consumers Directorate of the European Commission, focusing on digitalization of justice. Her research interest focuses on legal and ethical concerns regarding ICT, digitalization and AI governance and policy.

1 Titsias, D. (2021) 'The digitalisation of judicial cooperation in the EU and its impact on the Greek judicial system', *Elliniki Dikaiosyni*, vol. 62(5), pp. 1336.

At the same time, the access to legislation and case-law via the internet and the electronic filing and service, the use of videoconferencing along with the prospect of predictive justice constitute some further matters which need to be taken into consideration before being vastly applied[2]. The application of algorithms for profiling under the auspices of criminal justice in particular, although forbidden in the EU, is common outside its borders and should be considered even as a far-fetched scenario[3].

2. Online dispute resolution

Online dispute resolution may occur in different forms, each of which should be examined and confronted separately. Surely, the complete submission of the judicial activity to any kind of automated processes stays out of question since it is inconsistent with the right to a fair trial, to access to justice and to adversarial hearing. It goes without saying that the access to justice could be tremendously improved by a solution that combines the automation promised by new technologies with the principle of the natural judge, creating a hybrid model of dispute resolution. Such a development would enable a more effective and fast distribution of tasks inside the court, reduction of the operating costs and enforcement of alternative dispute resolution methods.

Despite the positive aspects, the gradual dominance of a dispute resolution system with hybrid characteristics would promote the constant use of new technologies in order to solve problems that do not necessarily and in principle fall under their scope. In that case justice would lose its balancing dynamic between the contradictory parties and end up as a simple digital settlement[4].

As a result, the establishment of certain guarantees is a necessity. First of all, the potentiality of inaccuracy in such proceedings run entirely by algorithms operating according to a convincing model of cognition based on artificial intelligence or machine learning technologies should not be excluded. Such a model, though, has not been applied in social sciences, like the one of law or justice and an effort of constructing it would eliminate the human factor and individualization, core elements of a judgement. It is true that in various cases of small claims online resolution methods have been applied. Such a choice should follow the explicit consent of all the involved parties, who would be capable of comprehending that they are subject to an alternative dispute resolution procedure, differentiating from the traditional one before the court.

---

2 Aletras, N., Tsarapatsanis, N., Preoţiuc-Pietro, D., Lampos, V. (2016) 'Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective', *PeerJ Computer Science*, vol. 93(2).

3 Završnik, A. (2021) 'Algorithmic justice: Algorithms and big data in criminal justice settings', *European Journal of Criminology*, vol. 18(5), pp. 623-642.

4 Morozov, E. (2014) *Pour tout resoudre cliquez ici*, FYP.

Otherwise, they would end up appearing without their knowledge before a private judicial institution, even if it appears as a public one[5], instead of the expected tribunal, with all the legal uncertainty that such a phenomenon comes with. Clear and comprehensive information should from the beginning be provided to all parties, so as the latter would definitely be aware whether their dispute would be resolved in a fully automated manner or with the participation of a mediator or an arbitrator and choose consciously to be part to or to abstain from it. Additionally, the possibility of challenging the decision before a real court should be at hand and communicated in advance to the parties. The verdicts should be subject to further judicial review and not constitute the absolute end of the procedure. Thus, the right to an appeal to courts of second or further instance should be available in order to safeguard the process of judicial review. Apart from the verdict, the interested party should have access to and the capacity to challenge the scientific validity of the algorithm. The lack of the transparency of the functioning of the algorithm designed by private companies, the confidentiality and privacy concerns, the competition between the involved industries and intellectual property rights hinder more the spread of knowledge regarding the reasoning methodology those systems follow to reach a decision. For these reasons, the importance given to the aforementioned factors ought to be diminished and the inaccuracy potentiality to be highlighted so as for the parties to be able to invoke it in case of appeal or for the judges to take it into account while forming their judgement. A judge discerning from the algorithmically drawn conclusion should not be obliged to special reasoning or take greater responsibility. Those online dispute resolution mechanisms must facilitate judges' work rather than posing obstacles, submitting them to doubts as far as the validity of their judgement or their stance towards a dispute go.

The transparency and neutrality of those systems are desirable, but difficult to achieve. Public authorities along with the judiciary and the assistance of technical staff should pose those programs under constant supervision and evaluation to determine their effectiveness, their efficiency and to avoid unforeseen ramifications which would turn against the citizens and, indeed, themselves.

3. Electronic access to case-law and legislation

A first and foremost concern of the policy-makers in Europe and worldwide is the amelioration of the access of the citizens to every part and parcel of justice, especially to the legal framework under which it is administrated, namely legislation and case-law. Unfortunately, aside from some exceptions, the access to both requires open data policies, which are not fully implemented in many

5 Ferrié, S.M. (2018) 'Les algorithmes à l'épreuve du droit au procès équitable', *La Semaine Juridique édition générale*, vol. 11.

EU Member States, where one can obtain a full spectrum of national legislation and case-law only as a member, by paying a considerable sum of money to a privately run database.

Even if the release of legislation to openly to the public and the limitless accessibility to all databases has not met objections, the same does not apply for the open nature of databases including case-law. In many cases, the anonymization of a court decision is out of question, since the removal of specific facts would render the judgement pointless, so the competent authorities feel sufficient with pure pseudonymization.

At the same time, the number, the variety of the content of the information found in a judgement and the sensitive nature of the data[6], in combination with the high possibility of cross-referencing does not vanish entirely the danger of identification[7]. The mining and spread of this information could violate human dignity and allow profiling and discriminatory practices. With respect to personal data protection laws, it is necessary to limit to the minimum references inside the body of judgements that might if combined with other sources lead to identification of a person. If anonymization is impossible, a balancing between the right to access to information by the public and the right to privacy of the parties involved should proceed the publication of the judgement. The establishment of a mechanism through which the data subjects could turn to the operator of the database and exercises all their rights provided for in the GDPR would be another feasible solution.

Apart from the identification of the parties, the one of the members of the tribunal is also at stake. The reference to personal information of the judges aims to guarantee their objectivity and impartiality, since in that way they are recognizable and their lawful appointment and attribution of duties can be verified, while their compliance to fundamental procedural rules, among which the publicity and the collective operation can also be examined[8]. A development as such could, nevertheless, spoil the principle of the natural judge and their impartiality, taking into consideration that a cross- examination of their public and private information would be possible. On the one hand, the parties or their representatives might engage in unfair practices to achieve a more favorable outcome for their case. Under those circumstances, the danger of bribery or of the appearance of unlawful circuits within the court arises. On the other hand,

6 Kanelos, L. (2021) Εφαρμογές Τεχνητής Νοημοσύνης (στο δίκαιο & στη δικαστική πρακτική), *Athens: Nomiki Bibliothiki*, p. 192.

7 European Commission (1998). Green paper public sector information: a key resource for Europe - Green Paper on public sector information in the information society (COM (1998) 585 final).

8 ECHR. Vernes v. France, app no.30183/06, 20 January 2011; Pretto v. Italy, app no.7984/77, 8 December 2003; Kontalexis v. Greece, app no.590000/08, 31 May 2011, cl. 38; DMD GROUP, a.s. v. Slovakia, app no.19334/03, 5 October 2010, cl. 66; Miracle Europe KFT v. Hungary, app no.57774/13, 12 January 2016, cl. 58.

the cross-examination between the content of a case-law database and the judicial staff's data stemming from other sources will violate not only their privacy, but also their impartial judgement. For example, certain opinions expressed by them whether they are political or ideological of any other kind might lead to questioning of the validity of their conclusion and pose them and their verdicts under a direct or indirect control, which will be hazardous for their impartiality. Under the constant fear of criticism, judges might end up to a typical application of the law, free from their individualized reasoning, something that will hurt their independency, their discretion, the objectivity they are entitled to as representatives of justice and freedom of expression and as mindful members of the collective sphere. The application of the precautionary principle and the assessment of the appropriateness per case and instance before the official publication of the names of the members of the tribunal in accordance with the General Data Protection Regulation is considered as a positive counter-effect.

4. Electronic filing and service

Despite the fact that the already adopted electronic filing and service of documents system contributes to procedural economy, to relief of the congestion in the courts and to the facilitation of the professionals, it gives birth to some issues that should not be overlooked. Their mandatory electronic nature, which is not in all cases yet applied, has not been fully accepted. More specifically, the digital illiteracy of the general public and the legal professionals will pose difficulties to their access to procedures, marginalizing those who are incapable of keeping up with the technological evolution. Simultaneously, the leakage of personal data of parties or third persons and the threat of cybersecurity must be bore in mind. The absence of safe technical infrastructure, training and education of the judicial staff intensifies the already-existing considerations[9]. To counterweight those thoughts, the implementation of risk assessments during the design of those systems and the continuous supervision of their function by technicians, but also by those responsible for the protection of personal data from incompatible with the purpose of the processing uses. An a priori management of relevant risks with the application of the appropriate preventive measures at the stage of development (security by design) and beforehand, at the regulatory stage (security by default)[10] is highly recommended.

9 Greek Administrative Justice Association, Newsletter, 17 September 2019.

10 Ronsin, X., Lampos, V. (2018) 'In-depth study on the use of AI in judicial systems, notably AI applications processing judicial decisions and data', *Strasbourg: Documents and Publications Production Department (SPDP).*

5. Videoconferencing

Videoconferencing can be utilized for administrative purposes in the judicial sector and for the conduct of court proceedings and procedural actions of any form. Legitimacy issues arise concerning the latter use. Already, European Court of Human Rights with two of its judgements[11] has deemed videoconferencing in criminal cases compatible with the requirements of Article 6 of European Convention of Human Rights (ECHR) as long as this mechanism reinforces the acceleration of the procedure under the condition of safeguarding the rights of the parties and the confidentiality[12]. On the other side stands the French State Council, which considered the possibility to impose videoconferencing before criminal court to be against ECHR[13]. More precisely, it concluded that the right to fair trial has been violated since the imposition of videoconferencing is neither subject to any legal requirement, nor accompanied by any criterion, with the excuse of the confrontation of the pandemic not being able to justify the aforementioned choice. These two diverse opinions reflect perfectly the existing dilemma Although, videoconferencing means time and expenses saving, acceleration of the proceedings and access in cases in which physical presence is extremely difficult for the parties, it is also associated with flow of personal data, demand of basic digital skills of the citizens, lawyers and judges and availability of financial resources. A framework of technical options with deontological aspect should be at hand. Solutions like this one, with the application of information technologies such as the services with direct connection, distant hearings and videoconferencing, future features of digital justice, should respect fundamental rights and principles of fair trial[14]. The application of videoconferencing must be limited and the ultimate choice, one guaranteeing and not hindering the access to justice. Towards this direction, there is an urgent need for judicial training and the adoption of a legal framework foreseeing common standards for conducting hearings via videoconference, possibly based on existing CEPEJ guidelines. This would help overcoming differences between the rules applied in the different Member States and could enhance the conduct of remote hearings bringing judicial cooperation even closer.

---

11 ECtHR. Sakhnovski v. Russia, app no. 21272/03, 2 November 2010; Marcello Viola v. Italy, app no. 45106/04, 5 Οκτωβρίου 2006.

12 Ibid.

13 Conseil d'État. no. 440037. 5 March 2021.

14 Papapanagiotou, A., Zachou, X. (2021) Συστήματα τεχνητής νοημοσύνης στον τομέα της δικαιοσύνης. Τα ηθικά διλήμματα και τα όρια του ευρωπαϊκού χάρτη δεοντολογίας. Ψηφιακή Δικαιοσύνη: Σύγχρονες Προκλήσεις και Προβληματισμοί. ΕΣΔΙ.

6. Predictive justice

Uder the term of predictive justice one can understand the outcome of analyzing of a huge number of judgements with the use of modern technologies targeting at the prediction of the result of the disputes under examination. For its deployment, an open data policy of and accessibility to judicial verdicts or any kind of judgement, along with the possibility of storage and further use of third parties is essential. Among them, legal technology companies with expertise in case-law search engines and tend analysis are expected to play a leading role. Software produced by those entities stores the received data, classifies them and by observing some repetitive patterns inside them, tends to reach a conclusion on whether a case would succeed or fail before the court. The possibilities are formed based on statistical modeling of former judgements, with the use of methods belonging to the field of computer science, natural language processing and machine learning. In relation to justice, predictive justice systems are designed to be used by legal services, insurers, lawyers to predict the outcome of a legal dispute. Their usage is not limited in the sectors mentioned above, as they would assist judges to reach their decision[15].

Legal uncertainty, namely the risk of the acceptance or the rejection of a claim strengthens the desire for a quantification with the deployment of new technologies[16]. In this way, better counseling services could be provided to the interested parties related to the empirical and systematic evaluation of the possibilities of success or failure of a case, something traditionally attributed to legal professionals. The systems would lead the latter and their clients towards the conclusion of an agreement, to avoid a time-consuming and expensive trial or generally towards the most advantageous outcome. Therefore, the need for balancing between legal certainty due to the predictiveness and the creative interpretation of the courts is acknowledged[17].

Statistical models with more precise scales on the average awarded sum of money have been designed, are already used and are thought to assist at a vital rate the administration of justice. However, they should be restrained to a simple calculation, solely helping the judges, who can verify the result afterwards, and not replacing them. This statistical processing of electronic data can reveal the frequency of the use of certain group of words or terms, but is not capable of spotting the real reasons behind a judgement and proceed to a legal analysis.

15 Ibid.

16 Hartmann, K., Wenzelburger, G. (2021) 'Uncertainty, Risk and the Use of Algorithms in Policy Decisions: A Case Study on Criminal Justice in the USA', *Policy Sciences,* vol. 54, pp. 269-287.

17 ECtHR. Lupeni Greek Catholic Parish and Others v. Romania [GC], app no. 76943/11, 19 May 2015, cl. 116.

AI systems cannot till now individualize and subsume, so they are restricted to computational acts, classifications and groupings.

If a process like this is transferred from a simple calculation to the core of judicial work, it is probable that it produces a sense of urgency, a novel kind of normativity, which would operate additional to the law, limiting the discretion of the judges and in the long run allow the standardization of court decisions, which will no longer base on an individualized thinking on behalf of the judges but on a clearly statistical calculation, sterile and deprived of the human element, a true nightmare for the rule of law. Judges would not reach their judgements according to the law, but the case-law trends, stemming from statistics gathered and processed by a digital tool[18]. The rule of the majority of court decisions would be standardized. Any declination would cause more issues which might entail the reasoning by the discerning party[19].

On the contrary, judges' rationale depends on the evaluation and interpretation of those crucial and proven facts of a case, of the applied legislation, the meaning of those remains unclear and their own view upon those. The complexity of law due to its doubtful character would be replaced by a fragile, imposed and threatening for the judicial values certainty. Thus, since reality cannot always be statistically represented and the outcome cannot always be correctly foreseen, lies the risk that the results of predictive justice software would be set as prototypes, without any verification by the legal system and, in many cases, even against it.

These systems would be developed under the auspices of the private sector and would be offered, moreover, after the payment of a certain fee to the public one, triggering some more questions. The use of those applications would enhance the distortion of competition and the inequality of arms between the legal professionals who have access to them and those who do not. Furthermore, it would violate the right to an attorney, if one takes into consideration that a lawyer would not take on a case algorithmically proved to be unsuccessful.

Last but not least, predictive analyses might adopt discriminatory practices, if the data input is based on samples. This danger could appear from the start in the form of an impartial choice of a specific set of information instead of another and continue to lie during the entire judicial process According to its study, the University College London concluded that an automated intelligence model would be able to predict the outcome of a case with 79% accuracy in front of a specific court[20]. Distribution of justice cannot render a lucky game and

18 Završnik, A. (2021) 'Algorithmic justice: Algorithms and big data in criminal justice settings', *European Journal of Criminology*, vol. 18(5), pp. 623-642

19 Buat-Ménard, E., Giambiasi, P. (2017) 'La mémoire numérique des décisions judiciaires', *Dalloz Actualite*.

20 Aletras, N., Tsarapatsanis, N., Preoţiuc-Pietro, D., Lampos, V. (2016) 'Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective',

should not rely upon forecast. Discrimination, normalization of the majority and its prevalence against an in concreto and ad hoc balancing and judgement, rejection of discretion and the limits posed to the principles of independence and the rights to an attorney and to access to justice constitute some of the ramifications of the domination of predictive justice systems inside and outside of the court. Their usage only for the resolution of specific disputes with rather calculatory nature would be seen as a possibility. Even that one would call upon consultation of the involved parties, continuous controls and constant assessment, in order for the judgements to correspond to legislative amendments and jurisprudence conversions.

7. Profiling

Algorithms are tools of logic and resolution[21]. Adherent to automation and the systems adapted to it is the possibility of false exclusions, even of discrimination[22]. It is generally accepted that profiling and classification allow different and discriminatory practices. The behavioral past of a person which forms their profile can affect a judicial judgement, but also the person's future as a whole. While human decisions are based among others on values and thoughts, not the same applies for automated decisions. Instead, when automated decisions are reached with the use of biometrics, misclassification or stigmatization are possible, placing automatically a person under a group like the one of a terrorist or a criminal. Such a classification may cause the reversal of the presumption of innocence. The accused will be seen as guilty until he is proved to be innocent and not vice versa. It is extremely difficult to get rid of this stigma, taking into account that it remains stored in a database[23]. The prospect of it is reinforced by the use of facial recognition technologies, which at the altar of national security tend to identify persons without their knowledge and consent, giving flesh and bone to a self-fulfilling prophecy[24]. The Dutch Data Protection Authority had already warned of a "digital fate"[25], the risk that one might not escape from the digital profile created for them. When profiling and risk assessment methods

---

PeerJ Computer Science, vol. 93(2).

21 Harkens, A. (2018) 'The Ghost in the Legal Machine: Algorithmic Governmentality, Economy, and the Practice of Law', *Journal of Information, Communication and Ethics in Society*, vol. 16(1).

22 Zarsky, T. (2013) 'Transparent Predictions', *University of Illinois Law Review,* vol. 4, p. 1560.

23 Sutrop, M. (2010) 'Ethical issues in governing biometric technologies. Ethics and policy of biometrics', *Lecture Notes in Computer Science*, edit A. Kumar & D. Zhang, vol. 6005. pp. 102-114.

24 Keats Citron, D., Pasquale, F.A. (2014) 'The scored society: Due process for automated predictions', *Washington Law Review*, vol. 89, p.1.

25 Kohnstamm, J. (2014) 'Digitale predestinatie Speech big data', *Nationale Denktank Expertforum*. Available at: https://autoriteitpersoonsgegevens.nl/uploads/imported/speech_big_data_nationale_denktank_versie_3_okt_2014_website.pdf.

are mobilized in the field of security and justice, mainly for the detection of criminals, fugitives or terrorists, presumption of innocence would falter. Even if data regarding a person are gathered normally, after they are placed under the regime of the suspect, big data and risk profiling tend to precede, overturning this order[26]. Those tools are expected to reproduce unfounded already-existing inequalities and to legitimize policies that are inconsistent with the law. At the same time, the main destination of the criminal justice tool, which is rehabilitation of the offender is overridden. The role of the judge and the experts at the individualization of the sentence after the examination of the personality of the offender is also undermined.

8. Conclusion

To sum up, a common approach seems to be hindered by the diverse national levels of digitalization across the EU, the fragmented regulatory framework and the different perceptions of technology as a whole. As a result, there are doubts whether overregulation is the appropriate answer to the urgent need of efficiency, transparency and rapidity in the delivery of justice. Concerns have been expressed regarding the impact of digitalisation procedures and penetration of new technologies into the judicial field to fundamental rights, EU values and the notion of freedom and democracy in general. Risks are posed to personal data, cybersecurity integrity and rights of the parties of court proceedings. Moreover, the possibility of deployment of biased algorithms in justice, of crowding- out of the judiciary and their impartial and independent thinking and reasoning, of their replacement with automated means reaching a decision based on statistics rather than an individualized assessment seems menacing for the whole European acquis[27].

A system combining the legal and the digital element, since both of them are constantly developing following the pace and the disputes of society should at any time adapt, without meaning that they should undergo alternations. Thus, it means that a system as such, a system "viator" cannot undergo total, disarming, exclusive, unique regulation, but should follow the line of the changes that have already taken place, are happening right now or are going to occur in the future. The inability to grasp this floating nature by the entities responsible for digital policy, governance or all those participating in the administration of justice justifies the unending law-making as a hopeless attempt to establish normativity and deontology. The shifts on technological and social level present us with

26 Royakkers, L., Timmer, J., Kool, L., van Est, R. (2018) 'Societal and ethical issues of digitization', *Ethics Inf Technol*, vol. 20. pp. 127-142.

27 Keats Citron, D., Pasquale, F.A. (2014) 'The scored society: Due process for automated predictions', *Washington Law Review*, vol. 89.

new challenges and ask for new, sustainable, balancing and acceptable by all interested parties' solutions.

Apparently, a few opportunities do exist to combat those negative aspects of digitalisation. The attribution of a more human-centric orientation to the relevant judicial framework, the provision and application of accountability mechanisms, the launch of public consultations[28], ensuring that the skeptical voices are heard could be the solutions to the aforementioned challenges. The combination of ethics and regulatory measures that do not sacrifice the human factor for the shake of a technological deterministic view that tends to praise speed and effectiveness sounds ideal[29].

28 Zarsky, T. (2013) 'Transparent Predictions', *University of Illinois Law Review,* vol. 4.

29 Dymitruk, M. (2019) 'Ethical Artificial Intelligence in judiciary', *Jusletter IT 21*.

Chapter XXII
Fairness By Design: A Value-Sensitive Approach to Exploring the Fairness Principle in the GDPR in the Context of Children's Interaction With AI Systems

by Ayça Atabey*

INDEX: 1. Fairness in data protection law. – 2. Relevance of Value Sensitive Design (VSD) in HCI. – 3. Fairness and meeting the expectations of children. – 4. The way forward: operationalizing fairness through the use of VSD.

1. Fairness in data protection law

The fairness principle in the GDPR should be read in accordance with the principle of best interests of the child in the UNCRC. The UK Age-Appropriate Design Code (AADC) is an excellent example that links data protection law with children's rights. It is particularly relevant in the context of the fairness principle since the AADC is underpinned by the fairness principle set out in Article 5(1)(a) of the GDPR, stating personal data shall be "processed lawfully, fairly and in a transparent manner in relation to the data subject (lawfulness, fairness, transparency)". The AADC (also referred to as the Children's code) is introduced by the UK Information Commissioner's Office (ICO) and contains 15 standards that online services need to follow and helps them comply with their obligations under data protection law to protect children's data online[1]. Importantly, the fairness principle in the GDPR also lies at the heart of all 15 standards of the AADC, particularly the best interests of the child standard (Standard 1). To comply with the fairness principle, among other requirements, companies also need to consider questions about balancing competing rights (e.g., rights to education and privacy when children interact with educational AI) and interests (those of the EdTech company and children). Yet, translating an abstract concept such as "fairness" (which has no definition in the GDPR) into

* Research Associate at Edinburgh University, she takes a look at the concept of 'fairness' in relation to children in the digital environment, explaining how crucial it is and how the UK's Age Appropriate Design Code offers an example of good practice of promoting fairness beyond just preventing unfairness.

1 See ICO "About this code".

technical design of digital technologies is not that straightforward. Similarly, it is a challenging task for designers to manage this balancing exercise, reflect it in the design of technologies, and ensure children's best interests are prioritized when processing children's data.

Fairness is an overarching concept that bolsters other principles like transparency, lawfulness and data minimisation. Compliance with fairness also means data controllers to consider balancing rights and interests, promoting children's control and agency over their data, and ensure that all children are treated equally and not discriminated against, and equality laws are also considered to ensure that children's best interests are prioritised in accordance with the AADC. To comply with the law, organizations need to consider children's needs and meet their expectations to ensure that their data processing can be considered "fair". Moreover, data controllers are also required to embed the fairness principle into the design of technologies. However, given the abstract nature of "fairness" and lack of its definition in the law, compliance with fairness becomes far from an easily achievable task. Considering the complexities about the meaning of "fairness", the interdisciplinary nature of data protection, and experts from different backgrounds and stakeholder groups being involved in decisions about design and data practices children become subject to in their daily lives, having a unified approach to "fairness" becomes utmost important. This becomes particularly crucial especially in relation to artificial intelligence (AI) where fairness is already being discussed with different definitions and perspectives focusing mainly on non-discrimination law. For data protection fairness, online services must design in accordance with child best interests, meaning that right to non-discrimination is by no means sufficient to ensure fairness by design.

Children have individual rights under the data protection laws in the EU and the UK. As the ICO notes "children have the same rights as adults over their personal data. These include the rights to access their personal data; request rectification; object to processing and have their personal data erased". Exercising data subject rights is linked to "transparency" which closely intersects with the fairness principle. This is also related to organisations' obligation to communicate information clearly, in an accessible and child-friendly way, and to consider the diverse needs of children. In deciding what information should be given to children, the ICO notes: "In order for processing to be fair, there is the same need for transparency, as this gives an individual control and choice"[2]. Aligned with ICO's approach to fair processing, we can say that transparency and fairness play critical roles in ensuring children can exercise their rights, but transparency is only one aspect of fairness and is rather as an enabler for fairness. This is because exercising rights is possible when children know about

2 See ICO (2022) GDPR Guidance.

their rights. This also helps organizations meet their obligations to facilitate the exercise of individuals' rights.

Overall, fairness is a broad, complex notion and is context dependent. It has several connotations in data protection law: procedural fairness, fair balancing, preventing adverse effect (e.g., discrimination, harms), and good faith (bona fide)[3]. One commonality among these four nuances is that fairness aims to rebalance significant imbalances experienced by the data subjects[4].

Fairness principle requires the data controllers to avoid exploiting data subjects' factors of vulnerability. Accordingly, it can be argued that the strong link between vulnerability and fairness stems from the power imbalance between the controllers and data subjects. This is why understanding vulnerability (both processing-based vulnerability and effect-based vulnerability) might serve as a tool to help solidifying fairness. Vulnerability is also exacerbated between stakeholder power dynamics, and therefore stakeholders' involvement in exploring fairness and understanding their perceptions of what's fair as well as current data practices that are impacted by existing power dynamics is important (e.g., education setting, teachers, edtech companies providing products to schools (data controllers), children). Future vague uses of fairness in principle "name-dropping" without developing a true understanding of what "fairness" represents or entails in practice should be avoided. Clifford and Ausloos clarified the overarching role of fairness and underscored its importance in the GDPR and noted that there is a need to precise contours of the fairness principle[5]. EU policymakers have a particular responsibility to put in place more detailed rules that set out in more detail what "fairness" entail in practice to protect vulnerable groups rather than leaving its interpretation to the courts and the regulators. Clarification would contribute to one of the purposes of the GDPR: unification among Member States. Clarification would also benefit different stakeholders. Firstly, it will be beneficial for the private sector who could seek certainty about what is expected from them for compliance reasons, especially when they build tech to be used by vulnerable people. Secondly, it could further benefit data subjects as they would know what to expect from data controllers and data processors. Thirdly, fleshing out "fairness principle" to a certain extent might help regulators as they would know what to look for when controllers demonstrate compliance with the fairness principle. Moreover, clarification could enable "fairness" as a clearer and stronger principle for code-based regulation, which may enhance consistency and predictability, and build

3 Clifford, D., Ausloos, J. (2018) 'Data Protection and the Role of fairness', *CiTiP Working Paper 29/2017*, Australian National University.

4 Atabey, A., Scarff, R. (2023) 'The Fairness Principle: A Tool to Protect Childrens Rights in Their Interaction with Emotional AI in Educational Settings', *Global Privacy Law Review*, vol. 4(1).

5 Clifford, D., Ausloos, J. (2018) 'Data Protection and the Role of fairness', *CiTiP Working Paper 29/2017*, Australian National University.

trust in the system. Without clarifying what "fairness" entails in practice, it may be challenging to use "fairness" as a tool to help prevent unfair power imbalances between vulnerable individuals and data controllers in the context of design of technologies, particularly in compliance with data protection by design requirements in the GDPR.

2. Relevance of Value Sensitive Design (VSD) in Human Computer Interaction (HCI)

Value-Sensitive Design (VSD) emerged in the 1990s as the most reviewed approach for integrating values in design[6]. The need for engineers to look beyond technical concerns has been recognised widely[7]. VSD is a framework that addresses this need. Nissenbaum argues that scientists need to consider social, ethical, and political criteria, and equally, social scientists have to look beyond theory and consider intricate technical details and how they interact with values[8]. There are examples showing that when values are not considered in design processes, undesired, unethical, biased, unfair outcomes could take place[9]. Integrating values into technologies and addressing difficult questions around the tensions between values such as privacy and security could be significant to build responsible technologies. Poel highlights VSD assumes "the configuration of technology is not value-neutral"[10]. Friedman and Hendry explain that technology is an outcome of human imagination, and "all technologies to some degree reflect, and reciprocally affect, human values" and "ignoring values in the design process is not a responsible option" because of this relationship between technologies and humans[11]. Accordingly, VSD seems to be a logical framework as it recognizes the connection between human values and technologies and addresses the need to incorporate values in design processes to build responsible technologies.

Friedman et al. defines VSD as "a theoretically grounded approach to the design of technology that accounts for human values in a principled and comprehensive manner throughout the design process"[12]. To elaborate, VSD com-

6 Manders-Huits, N. (2011) 'What Values in Design? The Challenge of Incorporating Moral Values into Design', *Science and Engineering Ethics*, vol. 17, p. 271; Friedman, B., Hendry, D.G. *Value Sensitive Design*. The MIT Press.

7 Friedman, B., Hendry, D.G. (2019), *Value Sensitive Design*. The MIT Press, pp. 211-212.

8 Nissenbaum, H. (2005) 'Values in Technical Design' in Carl Mitcham', *Encyclopaedia of Science Technology and Ethics*, MacMillan.

9 Friedman, B. (1996) 'Value-Sensitive Design', *Interactions*, vol. 3(16), pp. 16-23.

10 van de Poel, I. (2009) 'Values in Engineering Design', *Philosophy of Technology and Engineering Sciences*, Elsevier BV.

11 Friedman, B., Hendry, D.G., (2019) *Value Sensitive Design*, The MIT Press, pp. 211-212.

12 Friedman, B., et al. (2013) *Value Sensitive Design and Information Systems*, in Doorn, N., et al. (eds) (2013). *Early engagement and new technologies: Opening up the laboratory* Springer, vol.

bines theory grounds, mechanisms, and methods for considering values in a principled and comprehensive way. VSD is carried out not just at the beginning of the design process but rather as the system is being developed. Van Den Hoven provides a broader definition of VSD and describes it as "a way of engaging ICT that aims at making moral values part of technological design, research and development. It assumes that human values, norms, moral considerations can be imparted to the things we make and us[e].." further stating VSD construes information technology as a significant force that could be utilized "to make the world a better place, especially when we take the trouble of reflecting on its ethical aspects in advance"[13]. In short, VSD is arguably focused on the end product and how different values can be put into the design to create a value-sensitive system that is expected to reflect values.

In the VSD context, a value is considered that attribute with which an individual or group considers as important[14]. Value can be defined as "what a person or group of people consider important in life"[15], so-called human values with ethical import[16]. Although this is not a definite list, these values can be explained as "centre on human well-being, human dignity, justice, welfare and human rights"[17]. VSD is also about establishing risks, benefits, and costs from a particular design approach. It is also about finding direct and indirect stakeholders and looking at value conflicts and their resolution in a specific context[18]. This approach could make VSD valuable for building inclusive and fairness-aware technologies in specific contexts. This is because fairness is context-dependent and considering these factors for specific cases might help us better understand how values could be reflected into responsible technologies by exploring different aspects and this can be a step towards better conceptualising and making fairness more tangible in a specific context. Accordingly, fairness can also be explored from this point of view as it is also a value which is translated into practice through the perceptions of different stakeholders in the digital world.

16(56), (1st edn).

13 van den Hoven, J. (2007) 'ICT and Value Sensitive Design', The Information Society: Innovation, Legitimacy, Ethics and Democracy'. In honor of Professor Jacques Berleur s.j, *Springer US*, vol. 233(67).

14 Umbrello, S. (2020) 'Imaginative Value Sensitive Design: Using Moral Imagination Theory to Inform Responsible Technology Design', *Science and Engineering Ethics*, vol. 26, pp. 575-579.

15 Friedman, B., Kahn, P.H., Borning, A. (2009) 'Value Sensitive Design and Information Systems', *The Handbook of Information and Computer Ethics* (John Wiley & Sons, Inc 2009).

16 Friedman, B., et al. (2013) *Value Sensitive Design and Information Systems*, in Doorn, N., et al. (eds) (2013) *Early engagement and new technologies: Opening up the laboratory*, Springer, vol. 16(56), (1st edn).

17 Friedman, B., H Kahn, P. (2006) *Human Values, Ethics, and Design*, in Sears, A., Jacko, J. (eds) (2006). *The Human Computer Interaction Handbook*, CRC Press, 2nd edn, p. 1178.

18 Davis, J., Nathan, L.P. (2013) 'Value Sensitive Design: Applications, Adaptations, and Critiques. Handbook of Ethics, Values, and Technological Design', *Springer Netherlands*.

VSD has tripartite constituents. These are conceptual, empirical, and technical investigations that are carried out in an "iterative" and "integrative" way to unpack how values are involved in a system[19]. These three constituents feed into each other, which turns into an iterative process of refinement of change based on insights of each of these three stages. The conceptual constituent is "considered the most theoretical among the tripartite investigations"[20]. Conceptual work looks into normative frameworks and explores abstractly at philosophical questions of establishing values, balancing competing values, potential tensions between values, and possible effects of the system[21]. However, conceptual constituent has received criticism for not having an adequate ground for determining moral values. Umbrello argues, "the current account of VSD is predicated on an insufficient account of what constitutes values and moral deliberation"[22].

Empirical work grounds analysis with specific examples to look "individuals, groups, or larger social systems that use or are affected by the technology" (e.g., speaking to people or groups). Accordingly, empirical constituent might play a crucial role in understanding how stakeholders prioritise, perceive values and how they want those values to be engineered into responsible technologies. However, in practice, some challenges might exist, especially for identifying stakeholders. VSD identifies indirect stakeholders as a critical category of "user"; however, "the broad potential for social transformations of digital technologies implies that increasingly, the circle of indirect stakeholders becomes ever wider"[23]. VSD is also about understanding of stakeholder needs and values concerning a particular technology, and has been previously used and researched in relation to the design of technologies used by children, for example regarding "software tools that balance online risks and opportunities for

19 For a detailed explanation of tripartite constituents see Umbrello, S. (2020) 'Imaginative Value Sensitive Design: Using Moral Imagination Theory to Inform Responsible Technology Design', *Science and Engineering Ethics*, vol. 26, p. 578.; Friedman, B. (2019) 'Value Sensitive Design: Shaping Technology with Moral Imagination', (Hendry, D.G., ed). *MIT Press*, p. 31.

20 Umbrello, S. (2020) 'Imaginative Value Sensitive Design: Using Moral Imagination Theory to Inform Responsible Technology Design', *Science and Engineering Ethics*, vol. 26, p. 578; Friedman, B., H Kahn, P. (2006) *Human Values, Ethics, and Design*, in Sears, A., Jacko, J. (eds) (2006) *The Human Computer Interaction Handbook*, CRC Press, 2nd edn, p. 1181. Cited in Urquhart, L. (2017) 'Towards User Centric Regulation: Exploring the Interface between Information Technology Law and Human Computer Interaction', *Nottingham University*, p. 62. Available at: http://eprints.nottingham.ac.uk/41787/.

21 See Urquhart, L. (2017) 'Towards User Centric Regulation: Exploring the Interface between Information Technology Law and Human Computer Interaction', *Nottingham University*, p. 62. Available at: http://eprints.nottingham.ac.uk/41787/.

22 Umbrello, S. (2020) 'Imaginative Value Sensitive Design: Using Moral Imagination Theory to Inform Responsible Technology Design', *Science and Engineering Ethics*, vol. 26, p. 579.

23 Grimpe, B., Hartswood, M., Jirotka, M. (2014) 'Towards a Closer Dialogue between Policy and Practice: Responsible Design in HCI', *Conference on Human Factors in Computing Systems – Proceedings. Association for Computing Machinery*.

young children"[24]. The value-sensitive approach can further be seen in various children's digital experiences related research (e.g. Child rights by design work by Digital Futures Commission)[25] and participatory design studies in various disciplines[26].

The technical stage examines the technology context such as the nature of the system, the way it works, and the long-term management strategy plan for specific technologies by adopting a holistic approach. Technical investigations provide "value suitabilities that follow from properties of the technology" and technical work focuses on the way "technological properties and underlying mechanisms support or hinder human values"[27].

It can be argued that the fact that VSD encapsulates these three stages and allows going through these different steps make it possible to answer critical questions such as how the system is going to impact upon security, privacy, autonomy, or any other values and a balance among values can be achieved. It allows to take it from that level and work through the consequences and different ways that tensions or concerns can be addressed. However, it is also crucial to note that when there is a gap in practice in one of VSD's constituents, this might result in problematic outcomes which shows that unless the above-mentioned criticisms are not effectively addressed in practice, VSD might not be used to help design responsible technologies that processes children's data fairly. VSD framework's iterative approach could create challenges in practice. As Jongsma and Jongepier explain, "the strict division of conceptual, empirical and technological investigations risks undermining its iterative character"[28].

These three considerations might shape how technology is developed, and therefore by taking different approaches and steps into account, helping technologies can be designed in a more desired way. This sequential approach seems relevant in designing for fair data processing. However, the above-mentioned

---

24 Nouwen M., et al. (2015) 'A Value Sensitive Design Approach to Parental Software for Young Children IDC '15', *Medford*, MA, USA.

25 Livingstone, S., Pothong, K. (2023) 'Child rights by design toolkit and developing guidance for innovators work', *Digital Futures Commission.*

26 Skovbjerg, H.M., Bekker, T., Barendregt, W. (2016) 'Being Explicit about Underlying Values, Assumptions and Views when Designing for Children in the IDC Community', *Proceedings of the 15th International Conference on Interaction Design and Children – IDC '16*, pp. 713-719. Available at: https://doi.org/10.1145/2930674.2932224; Nouwen M., Van Mechelen, M., Zaman, B. (2015) 'A value sensitive design approach to parental software for young children', *Proceedings of the 14th International Conference on Interaction Design and Children – IDC '15*, pp. 363–366; Elsayed-Ali, S., Bonsignore, E., Hernisa, K., Subramaniam, M. (2020) 'Designing for Children's Values: Conceptualizing Value-Sensitive Technologies with Children IDC '20 Extended Abstracts'. Available at: https://doi.org/10.1145/3397617.3397826

27 Friedman, B., Kahn, P.H., Borning, A. (2009) 'Value Sensitive Design and Information Systems', *The Handbook of Information and Computer Ethics* (John Wiley & Sons, Inc 2009).

28 Jongsma, K.R., Jongepier, F. (2020) 'Value-Sensitive Design and Global Digital Health', *Bulletin of the World Health Organization*, vol. 98, pp. 579-580.

concerns should be addressed in practice. It could be argued that child rights discourse could contribute to VSD framework being used to explore fairness by design by thinking of ways which address the conflicting nature of different rights, responsibilities, and value tensions since such an approach is common in law (e.g., trying to find a balance between rights such as right to education, freedom of thought and privacy in children's interaction with Emotion AI[29]). Urquhart highlights that "Law is not just a static box of rules and regulations that can be translated directly into a system. Instead, it is a set of often conflicting high-level values, sometimes instantiated in legislation, sometimes in case law"[30]. Like tension between values that might exist in VSD, in data protection law, finding the fine balance between rights and balancing interests between data controller and children (legitimate interests assessments) are quite common. Of relevance here is the best interests of the child, which is the fairness principle, which also requires balancing different rights among others. However, the technical implications of these when designing AI systems for children also in compliance with the AADC is yet to be explored. Legal processes for balancing such competing rights (in the education context, privacy and education) could help resolve and manage conflicts of values in design. However, since finding the right balance between human rights can be challenging even for the courts[31] for VSD to consider best interests of the child among several stakeholders, there may be challenges and big differences and gaps in its considerations in practice. As such, for fairness by design to work in practice in compliance with data protection fairness there is first a need to explore what children's best interests entails and this should for sure include the voices of children in the discussions.

3. Fairness and meeting the expectations of children

Despite its challenges, VSD is a useful framework for reflecting values into technologies and designing for different vales of diverse stakeholders coming from different backgrounds and cultures, which also can have an effect on fairness and data privacy expectations. Designing for compliance with fairness in data protection would also require an inclusive but unified approach to ensure

29 Atabey, A., Scarff, R. (2023) 'The Fairness Principle: A Tool to Protect Childrens Rights in Their Interaction with Emotional AI in Educational Settings', *Global Privacy Law Review*, vol. 4(1).

30 Urquhart, L. (2017) 'Towards User Centric Regulation: Exploring the Interface between Information Technology Law and Human Computer Interaction', *Nottingham University*, p. 51. Available at: http://eprints.nottingham.ac.uk/41787/.

31 Greer, S. (2004) '"Balancing" and the European Court of Human Rights: A Contribution to the Habermas Alexy Debate', *Cambridge law journal*, vol. 63, pp. 412-434.; Christoffersen, J. (2009) *Fair Balance: Proportionality, Subsidiarity and Primarity in the European Convention on Human Rights*, International Studies in Human Rights, vol. 99.

that all children's data are processed fairly (different groups of children exist and their needs are expectations can be different) but also that responsible technologies cater for a universal approach to fairness that puts children's best interests at heart (universal principle in UNCRC). The above discussed practical challenges and concerns that are mostly related to VSD's iterative nature, making it difficult to ensure that VSD can help building responsible technologies that comply with fairness rules. Yet the VSD approach can inform exploration of fairness principle and its implications for design of AI systems for children to an extent especially by involving key stakeholders, most importantly, children themselves in the design processes while also questioning and exploring technical achievability of what the law says and how fairness is currently conceptualized by legal experts and law itself can work in practice and what would be needed by technical people to make it work in practice in children's best interests. This is particularly important given that data protection by design requires embedding privacy and data protection principles (including the fairness principle) into the design of data processing and business practices. The fairness principle and "data protection by design" in GDPR have critical roles in designing a fair and inclusive digital world that children deserve, where their expectations are considered. As the ICO notes, fairness is also about handling individuals' data "in ways they would reasonably expect". Therefore, to process data fairly, it is necessary to consider how it affects children and their interests more generally. This includes taking into account the fact that different groups of children can have different expectations and needs. Adopting this inclusive approach will help organizations comply with the law and promote accessibility and inclusiveness in the design processes – crucial to building an empowering and rights-respecting digital world where all children can benefit from digital technologies fairly and equally[32].

4. The way forward: operationalizing fairness through the use of VSD

Fairness lacks a specific definition under the data protection law. Despite being a core element that lies at the heart of the legal framework, it is difficult to understand the precise meaning of fairness principle. Due to its abstract and broad nature, its practical implications are also vague. Currently, there is limited regulatory guidance that deal with the fairness principle. The already existing research mainly focuses on the ethical aspect of "fairness", non-discrimination aspect mostly in the AI context, its supporting role to "transparency" and "lawfulness" and other rules, and intersection with consumer protection and competition law. However, these discussions seem to shed little light on the

32 Adapted from Digital Futures Commission blog post by Atabey (2022).

practical implications of "fairness principle" and its role in protecting vulnerable groups in practice/reality. As there is a lack of a precise understanding of what the fairness principle entails in practice, demonstrating compliance with it (for accountability purposes) could be challenging. Accordingly, how "fairness" is perceived by controllers and is applied in practice needs to be understood and the VSD method can be a crucial tool to realize what fairness can entail when designing digital technologies used by children.

Notably, participatory design in HCI and multistakeholder approaches in VSD can help us to further conceptualize and build what fairness by design can look like, especially given that understanding the expectations of children and ensuring children's data is processed fairly has become more relevant in today's world, where children face risks in their interaction with technologies. The core data protection principle of fairness provides an opportunity to understand children's expectations and needs when designing value-sensitive and child-rights respecting digital technologies. On the other hand, the VSD can contribute to the understanding of stakeholder needs and values concerning building "fair" data and design practices in that fairness doesn't only stay as an abstract concept in law but can truly be translated into practice in application of the law and this would also help address the potential gaps between what the law says and what is technically possible while furthering the exploration of what fairness by design can look like informed by the voices of different stakeholders' values and interests and needs balance risks and opportunities for children interacting with AI systems.

Overall, the VSD in HCI is a promising way towards exploring fairness from a nuanced/inclusive but also a holistic and universal/unified approach to protecting children's data[33]. This approach would also align with the AADC approaches to data processing (universal, because data and design practices must be children's best interests aligned with UNCRC principle BUT also vulnerability-aware, for example, by taking into account different vulnerable groups of children in accordance with equality legislation). The practical considerations of fairness and AADC standards (underpinned by the fairness principle in the GDPR) can also be further explored and discussed through addressing stakeholder needs and values concerning AI systems that can prioritize children's best interests, and help making fairness an inclusive, actionable and useful tool to be used in empowering children.

33 By involving different values informed by cultures, background differences, vulnerabilities of individuals, lives experiences etc.

Chapter XXIII
About the Need for Regulation Central Bank Digital Currency: Potential Monetary Legal Basis and Challenges

by Marko Dimitrijević*

INDEX: 1. Technological revolution and monetary/legal innovations. – 2. Digital money of central bank: classification and taxonomy. – 3. ECB monetary sovereignty and digital currencies.

1. Technological revolution and monetary/legal innovations

In the circumstances of accelerated technical-technological progress and the currents of the digital economy, data and information are basic goods that can be considered raw materials for processing from which new value can be generated. What's more, simple pieces of data that are processed for specific purposes in conjunction with other information represent a good basis for building models with predictive answers. These possibilities have led to the emergence of new applications for business models that mark the beginning of the era of (digital) capitalism, as it is popularly called recently[1]. Although the influence of

* Associate Professor (PhD) at the Faculty of Law University of Niš (The Republic of Serbia) and Jean Monnet Module for European Monetary Law – MONELA (2020-2023) academic coordinator. His fields of teaching include European and International Monetary Law, Economic Policy for Lawyers, International Financial Relations and Law of Economic System of the European Union. He is an observer of the European Law Institute University of Vienna (ELI), international research fellow at the Information Society Law Center (ISLC) at Cesare Beccaria Department of Law University of Milan (Italy) for 2022/2023 and member of Serbian Fiscal Society (branch of IFA Rotterdam). He is involved in various projects about the protection of human rights and Serbian law harmonization with EU Law. He is an author of scientific monograph of national importance entitled "International Monetary Law Institutions", "Law of the European Central Bank" and many scientific articles dealing with the issue of contemporary monetary law, which has been published in domestic and international scientific journals and presented in numerous international conferences. He developed scientific research in the field of the monetary law at Saarland University, European Institute (EU Cluster of Excellence for International and European Law), Saarbrücken (Germany), Max Planck Institute for Tax Law and Public Finances, Munich (Germany), Max Planck Institute for International, European and Regional Procedural Law (Luxembourg),

information technologies on the development of society has positive effects, as the capacity of individuals to exercise rights and freedoms increases, it has simultaneously led to new constitutional challenges. The possibilities provided by the so-called algorithmic technologies clash with their worrisome opacity and lack of responsibility, which is referred to in the literature as the "phenomenon of logocracy" as a new model in the management of society and economy.

The introduction of information technologies has largely changed the established understanding of the way monetary finance is carried out in practice. When we emphasize this, we do not mean the distinction between materialized and dematerialized money, i.e. demands of citizens towards credit agencies and intermediaries, but to different forms of electronic money. This becomes particularly noticeable with the widespread use of the so-called digital currency that takes place outside the framework of existing monetary legislation and without the control function of the central bank and the financial mediation of commercial banks. The monetary legislator will have numerous challenges in the way of legalizing such a method of payment and ensuring safety and legal protection, given that in various economic studies it is especially emphasized that with the expansion of its use, there will be an increase in the degree of systemic risk within the financial system (observed in a global context) due to the international character of electronic money[2]. Also, we fully agree with the so-called "Friedman's dilemma" which speaks of the fact that even in circumstances where electronic money would become dominant in the financial market, the central bank cannot lose its regulatory role, but can only speak of adjusting the techniques it uses when targeting inflation following institutional changes in the money market. Cryptocurrency, as currently the most represented form of digital currency, has certain inconsistencies in its use that open up space for numerous cases of abuse that may be left without adequate judicial protection, taking into account the absence of firm monetary law regulation, but the fact that citizens' participation in such transactions is based on voluntariness, conscious risk-taking, and space that is not under the "watchful eye" of the central bank. The frequency of such transactions in economic traffic, the acceptance of cryptocurrency as a legitimate means of payment by large multinational companies, and the trust that a considerable number of monetary users show in it emphasize the need for a certain type of legal regulation to preserve monetary stability and protect consumer rights. It is unequivocally

---

Max Planck Institute for International and Comparative Public Law in Heidelberg (Germany) and Institute for Comparative Law at Lausanne (Switzerland).

1 Micklitz, H.W. et al. (eds) (2022) *Constitutional Challenges in the Algorithmic Society*, Cambridge University Press.

2 Dimitrijević, M. (2018) 'Electronic Money in the Contemporary Monetary Law', *TCOP,* vol. 81; Dimitrijević, M. (2023) *Law of the European Central Bank*, Center for Publication, Niš Faculty of Law.

clear that cryptocurrencies represent a form of monetary innovation that in the broadest context can be seen as new instruments, means, and procedures that realize monetary preferences. It is obvious that the history of monetary and financial innovations is still not nearly written and that, in addition to the mentioned technical-technological process, it is influenced by numerous other factors such as inelastic demand for existing money, financial problems of governments, inter-jurisdictional competition in the field of monetary and fiscal policy, as and legal gaps in existing legal solutions that are used contrary to the general intention of the legislators[3]. Monetary innovations can achieve their full purpose provided that they are recognized as such and based on domestic monetary law regulations because legal recognition guarantees the preservation of the principle of already acquired rights and continuity and predictability in the application of monetary norms that inform citizens promptly of all the consequences of their monetary choices. It is also interesting that in addition to extensive studies on the need to introduce central bank digital money, IMF studies show that through the comprehensive application of linguistic, dogmatic, axiological, historical, and other dedicated methods of interpretation, an insignificant number of central bank laws today offer a potentially firm and sufficiently strong the legal basis for the digitization of the monopoly position of the central bank over the legal tender for determining the monetary unit.

Digital currencies in the broadest sense of the word represent a set of operations based on the application of algorithmic technologies that enable the conversion of processes from the real world to the analog world (bases) while adding some new properties to the existing conventional means of payment. Establishing a reliable mechanism for validating digital transactions is a big challenge for the monetary legislator because it creates confidence in the legitimacy of the actions of the contracting parties and the durability of the transaction. However, the mechanism as such does not only have legal features but in terms of its nature is multi-layered and must be the result of a multidisciplinary approach because it is based on asymmetric encryption methods and sophisticated decentralized redundant storage that cannot be compared to anything similar in the real (analog) world. Although currently cryptocurrencies are seen as an alternative money, we must note that monetary history was very rich in such examples. Conventional money as we know it today has gone a long way from the exchange of goods, gold, and silver, through the appearance of paper money, to the use of electronic money, the PayPal system, and cryptocurrency. Alternative money, unlike classic money, does not have all four well-known economic functions (accounting tool, means of payment, means of exchange, value preservation measure), but these partial functions cannot be ignored, because over

3 Bernholz, P., Vaubel, R. (2014) *Explaining Monetary and Financial Innovation: A Historical Analysis*, Springer, Munich.

time they can develop and compensate for their shortcomings. Unlike the first generation of cryptocurrencies, which was the subject of great interest, both in monetary law academia and in practice, the second generation of cryptocurrencies is still not the subject of extensive polemics. One of the reasons for the smaller number of works on the mentioned issue stems from the fact that these currencies have been in circulation for only a few years and that there are still no legally defined and uniform understandings of their concept and functions. The emergence of the second generation of cryptocurrencies (so-called stable currencies) is motivated by the intention of correcting the shortcomings of the first generation of cryptocurrencies, which concern the uncertainty and risk of their use in payment transactions. Accelerated technical-technological development has to a large extent, in a certain way, "challenged" the existing *status quo* of classical and to some extent contemporary monetary legal thoughts about how money is legally defined and who can be found in the role of the issuer of money, thereby somewhat shaking the centuries-old awareness of the nature and functions of money in modern society. Circumstances that have influenced the rethinking of already acquired habits include the emergence of reduced use of cash in circulation, the emergence of "sharing" technology that has enabled the creation of cryptocurrencies, announcements by leading technology and other multinational companies to issue private cryptocurrencies, as well as circumstances related to the global economic and financial crisis and pandemic that indicate certain advantages of alternative money.

2. Digital money of central bank: classification and taxonomy

When we talk about the digital money of the central bank, it should be emphasized that there is no universally accepted definition, but it most often means "any form of lump-sum responsibility of the central bank that is available to all economic agents without special conditions"[4]. Forms of central bank digital currency can occur as mandatory deposits of commercial banks with the central bank (discount rate policy), and hybrid prepaid cards or so-called "mobile wallets" that contain cryptocurrency. Both forms of digital currency would be based on the use of a special sharing technology that the central bank must possess as a prerequisite for issuing digital money. The basic *ratio* of the introduction of the new currency is initiated by the high demand and "popularity" of private digital money, with the fact that all risks related to its use with the digital money of the central bank are avoided because public (state) monetary management is behind it. The issuance of electronic money by the central bank

4 Meaning, J., et al. (2018) *Broadening Narrow Money: Monetary Policy with a Central Bank Digital Currency*. Bank Of England, SWP No. 724.

can have a significant psychological effect, which is reflected in the strengthening of the citizens' primordial connection with monetary sovereignty and the concept of monetary policy, which as a public policy should serve all citizens because we must not forget that monetary stability is a public good. From this good, we can also derive a special category of economic rights of citizens, such as the right to a stable domestic currency, which in modern technological conditions now also refers to the digital currency format, where only the state (central bank) with its apparatus and powers can provide and guarantee full protection. One of the ways to eliminate the risk could potentially be a solution to provide access to the general public indirectly through the mechanism of the so-called full reserve bank. At the basis of this mechanism is the idea that by determining the legal basis in the law on the operation of the central bank, it will be possible for new subjects to be able to hold reserves with the central bank, which would avoid "dilution" of the already existing functions of the central bank and avoid direct communication between the user of digital money and the issuer (which does not necessarily have to be bad, but in the current system of the central bank's jurisdiction it is not necessary). Adding new competencies to the central bank can be a complicated legislative procedure, bearing in mind the fact that the number of its tasks has increased significantly after the global financial crisis. The objectives for which the central bank is now responsible are very often in conflict with each other, which implies prior purification when choosing a specific objective as dominant at a certain moment in monetary history, but also raises the question of the responsibility of the central bank. Certainly, regular monitoring of events with the use of cryptocurrency can timely indicate the adoption of the necessary regulation, but this also requires good coordination with the competent bodies of the central bank. Cryptocurrencies are a kind of reflection of the times and lifestyle in which consumers want to reduce their financial, time, and psychological costs that arise during the classic satisfaction of purchase and sale performance, by the fact that they want to realize them from home, and not on the market itself. However, unlike electronic money, which falls under the *lex monetae* as the first and most important monetary prerogative of the state, cryptocurrencies are not regulated uniquely in monetary law. The reason for this is that they represent the result of technological innovations of individuals over which the central bank, as the guardian of monetary sovereignty, has no authority whatsoever, nor is the process of their creation in any way related to a legal tender for the determination of money, because it is based on the sharing economy and the use of mathematical algorithms that (currently) not found in monetary policy and legislation designed to protect and strengthen the influence of monetary sovereignty. This does not mean that the existing monetary and legal solutions are outdated and do not follow the course of time and all that technological revolutions bring with them, but on the contrary,

it indicates the real and logical need to invest further efforts by lawyers and economists for the optimal regulation of this phenomenon.

3. ECB monetary sovereignty and digital currencies

When it comes to a credible consideration of the need to issue a digital currency of central banks, it is important to point out the joint statement of the eight most important monetary institutions in the world, namely: Bank for International Settlements (BIS), Bank of Canada, ECB, Bank of Japan, Sveriges Riksbank, Switzerland national bank, the Bank of England and the Fed[5]. In the announcement, the central banks expressed the unequivocal need for the coexistence of digital currency with traditional *monnaie sripturelle*, if this is not how the component of flexibility and innovative character of modern fiscal systems is improved. It is important to note that in this statement the central banks take the position that the issuance of digital currency is not a goal *per se*, but is in the function of supporting all the goals of monetary policy where, given the different political and cultural preferences of central banks and different legal regulations, the future concrete (legislative) approach in the treatment and regulation of digital currency may show more or fewer differences[6]. Nevertheless, the importance of the document is reflected in the initial positioning of the main determinants of the digital currency, that is, the recognition of the need of the supreme subjects of international monetary law that the digital currency "exists" and that as such it will find its place in the monetary regulation. After all, it is not necessary to think about digital currency as a disloyal competitor of traditional money, but as a new form of means by which consumers want to satisfy their preferences in modern society. In support of this, Blueprint of digital euro was issued by the ECB, while a similar draft law on digital currencies was offered by the National Bank of Japan[7]. The release of the draft was preceded by an extensive public discussion, about whether the interested social and professional public would become familiar with its content since it is the first official publication of the ECB on the topic of digital money of the central bank. The Report was prepared by the Eurosystem expert group and approved by the ECB Executive Board, and it provides for the first time an official definition of the digital euro, looks at the potential effects of its issuance, considers various optimal design options and other issues of a technical and organizational nature. The ECB clearly emphasized that digital currency is not a substitute for classic money and that the process of issuing it must be based on synergy

5 BIS. Central Bank Digital Currencies: Foundational Principles and Core Features, Report No. 1, 2020.

6 Heinz, O. (2020) 'Introduction to the Panel on Central Bank Digital Currencies – in the Distant Future or Tomorrow?', *ESCB Legal Conference 2020*.

7 European Central Bank, Report on a Digital Euro, 2020.

with the industrial sector (which, among other things, promotes it). The report also sets out some of the foreseeable side effects of issuing a digital euro, which include the impact on the banking sector (including its role and funding) and financial stability, on the ECB's balance sheet, which is expected to be larger, but possibly more fragile. with issued digital currency; and, finally, to weaker digital currencies that could be more marginalized and (probably) lose importance. A significant issue regarding the regulation of the digital euro is also the question of the legal basis for its issuance, which may require changes to the founding acts, which may take a very long time. The process can be further prolonged due to the amendment of acts such as the directive on payment services, on electronic money, on the finality of settlements, against money laundering, or the general regulation on data protection, because the same would then have to be applied to the ECB as the issuer of digital currency. To ensure legal certainty, it is necessary to create conditions for the supervision of mediation and the supervision of the infrastructure involved in the distribution, holding, and transfer of the digital euro, as well as issues related to intellectual property and various other private law issues. In a rapidly changing world, the digital euro could support the goals of the Eurosystem by giving citizens access to a safe form of money. By issuing the digital euro, the Eurosystem would offer direct access to the money of the central bank to all citizens and companies, where the digital euro could enable citizens and companies to make payments with a simple, risk-free, and reliable digital means of payment accepted throughout the eurozone and thereby contribute to the preservation of monetary stability[8]. The analysis showed that the digital euro could be issued for the monetary support of the digitalization trend of the European economy and the preservation of the strategic independence of the European Union; in response to a significant decline in the role of cash as a means of payment; if there is significant potential for foreign central bank digital currencies or private digital payments to become widely used in the euro area; as a new channel of monetary policy transmission; to mitigate risks for the normal provision of payment services; to protect the international role of the euro in world financial flows, and to support the ecological approach in the field of monetary and business finance.

Considering the potential modalities of the digital euro, the ECB in the Report identifies two general types of digital euro, namely: offline digital euro and online digital euro. These types are compatible with each other and could be offered simultaneously to the extent that both satisfy the basic principles of the Eurosystem and meet the generally established requirements. Users could access the digital euro directly or through supervised (by the ECB) intermediaries. If users have direct access, the central bank would have to provide end-user-facing services such as customer identification and support, which are

8 Bindseil, U. (2020) 'Issuing a Digital Euro', *ESCB Legal Conference 2020*.

tasks currently not performed by the Eurosystem. If users would have access to the digital euro indirectly, that is, through intermediaries responsible for providing such services, user identification and support would be provided by the intermediaries themselves (and not the central bank). The offline digital euro could be used without the intervention of a third party and therefore should only be available through certain user devices, which can be distributed and/or financed through supervised intermediaries where it is a particular challenge to ensure protection against hacking. Using appropriate technical restrictions in the payment device could ensure anonymity and protection of personal data. The main characteristics of this form of digital euro are universality, lack of additional costs for the potential user, and ease of accessibility, as there is no need to provide an internet connection. The online form of the digital euro could be charged at a rate that varies over time, where the fee would be an important tool of monetary policy and at the same time limit (act as a deterrent) the switch from private money to the digital euro (although for this purpose the transmission of monetary policies). Its use would not be tied to a specific device and access to all digital euro services could be controlled by responsible parties (central bank and supervised private intermediaries) at any time. However, the online digital euro would exclude the possibility of anonymity for users, unlike the offline euro. It should be noted that any digital euro for offline use would need to be managed online at some point to add funds to the device (or withdraw funds). It is important to point out that the use of one form of digital euro does not exclude the other, which is very useful for application in practice because the disadvantages of one form are compensated by the advantages of another form of digital euro. The issues arising from the application of technological innovations in the field of ECB law concern the protection not only of monetary sovereignty and the operation of public policies but also of the central bank's ability to preserve and enforce collective choices in terms of monetary and financial stability while making maximum use of technologies in monetary management (but preventing the emergence of technological monopolies) and avoiding the adoption of partial, non-systemic solutions with limited interoperability and the promotion of financial inclusion.

A different concept in the process of legal regulation of the central bank's digital money is quite radical and instead of complementarity, it emphasizes the need to replace classic money with digital currencies stating that is the monetary scenario that surrounds us today very similar to the one that preceded the issuance of banknotes in the 13th century (which perhaps one day the chroniclers of monetary history will write about as a turning point in the history of the development of money)[9]. Although we agree that in the process of issuing digital currency of the central bank, a clear distinction should be made between what is

9 Papapaschalis, P. (2020) 'Retail Central Bank Digital Currency: A (Legal) Novelty?', *ESCB Legal Conference 2020*.

legal and what are other political and technological conditions that must be met in advance to provide meaning to the process, we also believe that simplifying the facts is not necessary either. and that it can be an aggravating circumstance on the path to optimal legal regulation of digital currency, and that the path that the monetary legislator should follow is the one that has already been successfully mastered when issuing electronic money. When it comes to the monetary and legal aspects of digital money, some IMF experts point out that the digital money of the central bank would not change anything significant in the axiology of the normative regulation of monetary relations. Their explanation of such a position is based on the fact that digital money does not represent a new currency unit, because digital money is only spoken of as a digitized form of traditional money, which means that it is only a "new design" and form of expression of already existing money, and not introducing a new currency unit. This implies that there are no changes in the way the norms of monetary law and the rights of the central bank are operationalized, because the only thing new in the structure of the law of central banks is that the central bank gets a new authority to simply issue a digital form of an officially denominated monetary unit, such as banknotes, coins, book money and bills of exchange[10]. Of course, by changing the monetary legislation, it would be possible to define a digital currency as a new currency unit of a specific monetary jurisdiction, but the prerequisite for such a change is that the legislation prescribes a reliable (digital) conversion mechanism for determining the exchange rate with the existing monetary unit of the country. This would practically mean that in the structure of the domestic monetary policy, a distinction would be made between the monetary regime of two parallel and completely equal monetary units, but in this connection, the question of the concrete benefits of such a complex solution arises. Taking into account all the mentioned difficulties and challenges, the results of the monetary projections in the EMU point to the conclusion that the central bank's digital money will probably not be fully legally regulated before 2026, because despite the increase in the demand for economic entities for payment solutions that use the central bank's sharing technique they still do not possess the necessary expertise to follow these trends in the real economy[11]. The reason for this prolongation is the concern of the central bank due to neglecting the intermediary function of the banking sector, which is their traditional role because the development of the economy is based on loans from commercial banks. Nevertheless, the presence of smart contracts increasingly relativizes this role of the banking sector, which requires an appropriate response from the central bank to prevent the gains made on that front. The prerequisite for enjoying the benefits from the introduction of digital money

10 Bossu, W. et al. (2020) *Legal Aspects of Central Bank Digital Currency: Central Bank and Monetary Law Considerations*, IMF LDWP 254/2020.

11 Heckel, M., Waldenberger, F. (2022) *The Future of Financial Systems in the Digital Age Perspectives from Europe and Japan*, Springer, Munich.

by the central bank is the full implementation of interoperability, i.e. the ability for synchronized joint work of various systems, techniques, or organizations now directed towards the monetary, state, and IT sectors. Interoperability as an ability of heterogeneous systems requires that these systems work as well as possible, so that information can be exchanged, that is, so that it is available to the user, without requiring additional operations (requirements) for communication between two systems, which in the field of digitization of the norms of central bank law, it has not yet been achieved to the required extent. Legislative initiatives that would accompany the introduction of central bank digital money also depend on the specific form of that money, more precisely, whether it would be based on the introduction of tokens or central bank assets. If that digital money expressed in tokens were to be legally equated with banknotes (to the extent that this is even possible), the conditions for it to receive the status of legal tender should be established beforehand. One option in this regard is to limit the status of legal tender to a closed circle of sophisticated entities (state, public services, large traders, and/or firms with authorized activities, such as banks). After that, it would be necessary to carefully analyze the private law classification of digital money based on tokens and whether it is justified to give it certain advantages (incentives) arising from the principles of private law in terms of promoting use (advantages and benefits of use in payment transactions should be regulated by new solutions to popularize the use of central bank digital money). In the end, the legalization of the digital money of the central bank would also mean the need to redefine the criminal legislation in the part that refers to the legal nature of criminal acts of cybercrime, but also of financial crime (in general). At this point, it is important to note that in the monetary law of the EU, the terms digital money and cryptocurrency are not synonymous, because the term digital currency also includes electronic money that is accepted as a legal tender and represents a government debt document that is stored in a microprocessor or special protected software that is regulated by the law on the work of the central bank and accepted as a legitimate means of payment. Hypothetically speaking, based on the solutions that exist in tax law when citizens with the permission of the competent state authority can collect certain information in the tax collection process (when registering payments and taxpayers), similar solutions, let's call them, "borrowed monetary prerogatives could be assigned to physical persons but with the control of the supreme monetary institution (central bank)". However, in the first case, it is only about the technical division of work and actions that have a mechanical character (which is justified by the effect of congestion and the impossibility of taking all timely actions for the collection and control of taxes), while in the second case, it would be a significant ideological and axiological issue when, why and under what conditions an individual on behalf of the state disposes of "full monetary legal status", which is inseparable from political sovereignty.

Chapter XXIV
A.I., Facial Recognition and Privacy Risk

by Nicolò Bottura*

INDEX: 1. Facial recognition and use of biometric data. – 2. Applications of biometric facial recognition in the world: a brief overview. – 3. Italy's position and European recent developments. – 4. The importance of Privacy. – 5. Profiling Risk and actions useful to mitigate it. – 6. Ideas and reflections.

1. Facial recognition and use of biometric data

Facial recognition is a long-standing challenge in the field of Artificial Intelligence (A.I). It will be increasingly widespread, with huge potential. Market estimates, in fact, predict that the biometrics sector will double its turnover from $33 billion in 2019 to $65.3 billion by 2024[1]. The above-mentioned technology belongs to the branch of deep learning and can be defined as "the automatic processing of digital images containing faces of individuals"[2]. This system allows a mathematical formula to function as a human brain, performing a match as a result of collecting a large amount of data[3]. It involves the automated extraction, digitalization, and comparison of spatial and geometric distribution

---

* Lawyer, privacy & compliance analyst at the Bonelli Erede Lombardi Pappalardo Law Firm in Milan. Data protection and artificial intelligence scholar, he is a member of ENIA (National Agency for Artificial Intelligence) since April 2024. Always passionate about new technologies, he combined his studies on AI with cinema by directing the first edition of a film festival focusing on the theme of artificial intelligence and sustainability.

1 Tundo, E. (2022) *Biometric Surveillance Boom: Here's What Gains in Europe*, Editoriale Domani. Article represents an anticipation of an inquiry by Presa Diretta entitled "Weapons of Mass Control" by Giulia Bosetti and Eleonora Tundo, aired on Rai 3 on 17 October 2022. The information and data here quoted comes from the market analysis published by Global Market Insight for the period 2017-2024 in Biometrics Market Size, Growth – Industry Share Report 2017-20234, Global Market Insight Inc., August 2017.

2 This definition comes from Working Party Article 29, Opinion 02/2012 on facial recognition in online and mobile services, 00727/12/EN, WP 192, 22 March 2022, 2.

3 Wiener, N. (2019) *Cybernetics or Control and Communication in the Animal and the Machine*, The MIT Press.

of facial features to identify individuals[4] with the purpose of detect, recognize, verify and understand characteristics of human faces[5].

Our face can, therefore, be used to unlock the smartphone; it can be detected by CCTV cameras to compare it to those on a list of suspects or at the entrance of buildings or as part of e-boarding procedures at airports.

The individual may be subjected to a facial recognition process for condescension, or simply without being aware of it[6]. Using a digital photograph of a subject's face, a contour map of the position of facial features is converted into a digital template, using an algorithm to compare an image of a face with one stored in a database, usually collected from a vast number of images that have been uploaded to social media sites. This means that such technology uses biometric data. These are data relating to the physical, physiological, or behavioral characteristics of a natural person, which allow or confirm the unique identification of that natural person, such as facial images or dactyloscopic data.

Article 10, par. 1, Directive 2016/680 (Law Enforcement Directive) define biometric data as special categories of personal data, as from this data we have specific information about the individual. They reveal important aspects such as ethnic and racial origins, religious and political group membership, sexual orientation, and health. These are personal data deserving of enhanced protection such that their processing is prohibited, except for some cases expressly provided for in article 9 GDPR. This is because of their relevance to the person and his experience and their suitability, in case of misuse, to expose the subject to discrimination[7]. Biometric data show us how our bodies are increasingly technological, in the sense that they are the object of a decomposition process where every aspect is collected, stored, and delivered to a machine with the aim of subjecting it to an algorithmic analysis process[8]. To this extent, it's therefore necessary to ask ourselves a question: the use of biometric data represents a risk for the rights and freedoms of individuals?

4 Smith, M., Miller, S. (2021) *The ethical application of biometric facial recognition technology*, AI & Society, Springer.

5 Merner, M., Ratha, N., Feris, R.S., Smith, J.R. (2019) 'Diversity in faces', *IBM Research A*.I. arXiv:1901.10436

6 Berle, I. (2020) 'Face Recognition Technology. Compulsory Visibility and Its Impact on Privacy and the Confidentiality of Personal Identifiable Images', *Springer*.

7 Resta, F., Pollicino, O. (2020) 'Face recognition and data protection beware of the point of no return', *Diritti Comparati*.

8 Paolucci, F. (2021) 'Facial recognition and fundamental rights: surveillance is a fair price to pay?', *MediaLaws*.

2. Applications of biometric facial recognition in the world: a brief overview

There have been significant applications in relation to biometric facial recognition in all the world over the past ten years, and systems continue to advance rapidly. Their use in association with passports at international airports has been well established for more than a decade and they continue to play an important role in border controls systems. In recent times, this technology has become important for law enforcement investigations and for private companies.

In Australia, was introduced a regulation in 2015, permitting the release of biometric drivers' license photographs to NSW Police as well as the Australian Federal Police for the purposes of investigation of relevant criminal activity.

In the United States, since 2020, the law enforcement agencies and some service private companies (i.e. NBA, Walmart, Bank of America) have been using a biometric facial recognition algorithm, developed by the Company Clearview AI, to search images on internet to identify suspects and for private security purposes. In China, instead, use of biometric facial recognition is the rule of "living in public". It is already used in some school to analyze student reactions to lessons or to identify individuals in public places via CCTV who are suspected of minor crimes, such as jaywalking or for shaming[9] citizens engaging in "uncivilised behavior". India uses iris scans as part of its national identity system. The so called *Aadhaar project*[10], in fact, is aimed at uniquely establishing the identity of citizens and it is preordered to the more effective allocation of subsidies, benefits and public services to the neediest and the poorest sectors of the society[11]. In the UK, it is used in the railway stations and in the airports to control passengers. In the European Union post Brexit, France could be the first country to implement a facial recognition system aimed to create a digital identity spread to all citizens, even if the CNIL immediately said the opposite, warning that a digital identity project based on a mandatory facial recognition device would make the consent invalid because it is not free[12]. As highlighted

9 British Broadcasting Corporation – BBC (2020). Pyjamas in public: Chinese city apologises for 'shaming' residents. Available at: https://www.bbc.com/news/world-asia-china-51188669.

10 The discipline brought by the Aadhaar Act no. 18/2016 was challenged before the Indian Supreme Court, which by a ruling of 26 September 2018 was without prejudice to the system, provided, inter alia, that such forms of detection and control were limited to the original purpose; see Formici, G., (2019). Recognition systems and biometric data: a new challenge for Legislators and Courts, *DPCE online*, vol. 39(2), p. 1113.

11 See Mobilio, G. (2021) *Face recognition technologies: risks to fundamental rights and regulatory challenges*, Editoriale Scientifica.

12 Regarding free consent, with the sentence no. 1901249 of 3 February 2020, the Administrative Court of Marseilles has recently been pronounced on the installation in a school of cameras that performed facial recognition on students, establishing that the installation of such

by the EU's line, expressed also in the A.I. ACT[13], the use of these systems is not always under control. There are many cases in which there is a reckless use of these technologies, a use that is turning them into a discriminatory tool.

Famous in this regard is the Spanish case of Mercadona supermarkets[14], which were equipped with a video surveillance system that analyzed the biometric data of the face of customers who entered to do shopping to find out if they had slopes with justice, in this case blocking the entrance and triggering an alarm to request the intervention of security personnel.

3. Italy's position and European recent developments

In Italy, in line with the EU's orientation, law no.205/2021, law converting the Decree no. 139/2021 ("Urgent provisions for access to cultural, sports and recreational activities, as well as for the organization of public administrations and in the field of personal data protection"), has provided the suspension until December 31, 2023[15] – extended until December 31, 2025 with decree no. 51/2023[16] – of the installation and use in public places of video surveillance systems with facial recognition technologies operating through the use of biometric data[17] as there is a several questions about the balance between different interests and principles, as that of proportionality have been raised.

The processing carried out by the competent authorities for the purpose of preventing and prosecuting criminal offences or the execution of criminal

---

devices represents a violation of article 9 GDPR as students could neither provide nor deny their consent.

13 The political agreement on the A.I. ACT, reached on 9 December 2023, followed on 21 January 2024 by the draft of the final text, is the first European law on A.I. Among the prohibited uses provided for by the rules (object of long discussion between the commissions) there is also the use of face recognition and biometric systems in real time, which are discriminatory, except in three defined cases: the search for victims of crimes and missing persons; certain threats to the life or physical security of persons or terrorist attack; the location and identification of the alleged perpetrators of a list of 16 offences contained in an Annex IIa.

14 With the procedure no. PS/00120/2021, the Spanish Data Protection Autority (AEPD) has sanctioned Mercadona supermarkets for 2,5 million euros for illicit use of a video surveillance system equipped with facial recognition as well as for failure to comply with article 9 GDPR.

15 Until the entry into force of a legislative discipline of the matter and in any case no later than 31 December 2023, article 9, cit. In these terms already the amendment 9.500, Senate of the Republic XVIII Legislature, Iter Dossier DDL S. 2409, 21/11/2021.

16 With decree no. 51/2023, named Urgent provisions on the administration of public bodies, legislative deadlines and social solidarity initiatives and approved by the Italian Senate on 28 June 2023, the government, in line with the current European orientation about AI, has established the extension of the moratorium on facial recognition systems.

17 This solution was suggested during the consultations that subsequently led to the drafting of the White Paper on Artificial Intelligence.

penalties is an exception. The control on the legality of this video surveillance systems, also implemented at local level, is currently left to the Data Protection Authority. Already in 2019, the Authority expressed itself in contrast to the use of an automatic system to search for the identity of a face present in an image within a real-time database (SARI Real-Time). Due to the fact it would be an automated processing on large scale capable of revealing sensitive data such as trade union and religious beliefs, which cannot rely on a valid legitimate legal basis[18].

With the Order of Injunction of 10th February 2022, the Italian SA fined the US-based company Clearview AI Euro 20 million after finding it applied what amounted to biometric monitoring techniques also to individuals on the Italian territory. As highlighted in the above-mentioned document, the processing of personal data carried out by the company was illegal because it was carried out in violation of articles 5, par. 1, lett. a), b) and e), 6, 9. 12, 13, 14, 15, 27 of the GDPR[19].

Personal data held by the company, including biometric and geolocation information, in fact, were processed unlawfully without an appropriate legal basis – since the legitimate interest of the US-based company does not qualify as such. Additionally, the company infringed several fundamental principles of the GDPR including transparency – as it failed to adequately inform users –, purpose limitation – as it processed users' data other than those for which they had been made available online –, and storage limitation – as it did not set out any data storage period. Still on facial recognition, in November 2022, although in many Italian cities have already been installed surveillance systems of this type, the Italian Data Protection Authority opened an investigation against the Municipality of Lecce and Arezzo, which have announced, respectively, the start of a system that involves the use of facial recognition technologies and the testing of smart glasses, able to detect traffic offences from the license plate number of a car and verify the validity of the driver's documents[20]. Moreover,

18 GPDP, Opinion on the Sari Real Time System –25 March 2021 (9575877), 25 March 2021. Coherently, the Proposed Regulation indicates biometric identification in real time as "particularly intrusive in the rights and freedoms of the concerned persons, to the extent that it may affect the private life of a large part of the population, evoke a felling of constant surveillance and indirectly dissuade the exercise of the freedom of assembly and other fundamental rights". 2021/0106 (COD) Proposal, Recital no. 28.

19 The violation of some of the same principles also emerged in a recent measure of the Data Protection Authority against the Municipality of Trento: the Authority, by order of injunction of 11 January 2024, has detected the violation of privacy because of the installation of cameras in public areas with built-in microphone to train artificial intelligence systems, imposing a fine of Euro 50.000.

20 As reported in the press release Video-surveillance: Authority stop facial recognition and smart glasses. The Authority opens an investigation against two municipalities of 14 November 2022, the Data Protection Authority has warned against the use of smart glasses that may involve – even indirectly – remote monitoring of the worker's activities and called for compliance with the guarantees provided for by the Privacy rules and the Workers' Statute.

with Order of Injunction no. 369 of 10.11.2022, the Authority fined a sports club (Sportitalia) Eur. 20.000 for unlawful processing of biometric data of employees, collected through a fingerprinting system without an appropriate legal basis[21].

These last three measures take over from the guidance expressed by Council of Europe in its recent publication, where a ban – already asked by the EDPS[22] – of some facial recognition applications is required "for the sole purpose of determining a person's skin colour, religious or other belief, sex, racial or ethnic origin, age, health or social status to be prohibited"[23].

The document, addressed to Governments, Public Administrations, developers of facial recognition systems and companies to protect Privacy and right to identity, states that facial recognition cannot be used in uncontrolled environments, for the sole purpose of gathering data that may lead to the identity of the concerned person[24].

Although the guidelines provide that the use of facial recognition systems should be subject to a necessity and proportionality check and state that consent cannot be the legal basis for the processing of data, on the other hand, they show that there are real risks to the rights and freedoms of individuals. In some areas (e.g., police investigations) the use of such technologies can make a contribution that is difficult to make otherwise. In other contexts - and above all if it is conceived in a purely facilitative way of activities that can be realized in other ways (e.g., to control the exit of students) – it can instead result in an unjustified limitation of individual rights. Moreover, the widespread use of these techniques in ordinary circumstances and for mere facilitation purposes risks inducing a collective underestimation of the impact of such technologies[25]. Geolocation and targeted data profiling represents the most significant consequences of such underestimation.

Aware of these risks, the European Commission, as part of the drafting of the White Paper on Artificial Intelligence, in order to maintain the balance between

---

21 The text of the document states that "Sportitalia had processed a particular category of personal data for ordinary management purposes in breach of the principle of minimization and proportionality". In this context, a fundamental principle was also reiterated by the Authority: the processing of biometric data in the workplace is allowed only if necessary to fulfil the obligations and exercise the rights of the employer provided for by the legislation and with adequate guarantees.

22 EDPS, Artificial Intelligence Act: a welcomed initiative, but ban on remote biometric identification in public space is necessary, 23 April 2021.

23 Consultative Committee of the Convention for the protection of individuals with regard to automatic processing of personal data, Convention 108: Guidelines on facial recognition, January 28, 2021 mentioned also by F. Paolucci, F. (2021) 'Facial recognition and fundamental rights: is surveillance a fair price to pay?', *Media Laws*.

24 Raffiotta, E., Baroni, M. (2022) 'Artificial intelligence, identification tools and identity protection', *BioLaw Journal*, vol. 1.

25 On this point, Resta, F., Pollicino, O. (2020) 'Face recognition and data protection beware of the point of no return', *Diritti Comparati*.

individual freedom (privacy) and security, has attempted in vain to introduce a *moratorium* on the use of facial recognition in public places, for a period of three to five years. Although the Guidelines on facial recognition of 2021 have not closed this gap, it seems there will be a reversal: article 5 of the Proposal for a Regulation laying down harmonized rules on Artificial Intelligence[26] classifies the use of real time remote biometric identification for law enforcement as to be prohibited, providing in any case a series of exceptions.

At present, however, it must be considered that the regulatory gap contributes to undermine Privacy[27] of the individuals.

4. The importance of Privacy

Antonello Soro, during a meeting held in Rome[28], remembers Stefano Rodotà, who during his Closing speech of the 26th International Conference on Privacy and Personal Data Protection, which took place in Wroclaw on 14th, 15th, 16th September 2004, declared Privacy as "protection of existential choices against public control and social stigmatization" or "as the request for social tools that protect us from the risk of being simplified, objectified and judged out of context"[29]. Recognized in our legal system as a right of personality only in 1975[30], Privacy can be defined first of all as a right that people have in relation to other persons, the state, the organizations with respect to the possession of information about themselves by other persons and by organizations (e.g., images stored in biometric database) or the observation of themselves by other persons (e.g., via surveillance systems). Second, the right to privacy is closely related to the more fundamental moral value of autonomy, that consists of a right to exclude organizations and other individuals both from personal information and from the private sphere[31]. These connotations have a common objective, such as respect for the person from intrusions into his private sphere,

26 The whole text of the Proposal is available at: https://digital-strategy.ec.europa.eu/en/library/proposal-regulation-laying-down-harmonised-rules-artificial-intelligence. The final text version of A.I. ACT is available at: https://data.consilium.europa.eu/doc/document/ST-5662-2024-INIT/en/pdf.

27 First, Privacy was defined as right "to be let alone" in Warren, S.D., L. Brandeis, L. (1980) 'The Right of Privacy', *Harvard Law Review*.

28 Meeting: "Towards a new privacy?" In memory of Stefano Rodotà – Rome, 6 October 2017, available at: https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/6937167.

29 Privacy, Freedom, Dignity: Closing speech of the 26th International Conference on Privacy and Personal Data Protection available at: https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/1049293

30 Cass. Civ., sez. I, 27 May 1975 no. 2129.

31 Smith, M., Miller, S. (2021) *The ethical application of biometric facial recognition technology*, AI & Society, Springer; this definition is also mentioned in Kleinig, J. (2011) *Security and Privacy*, Anu press, Canberra. Available at: https://www.jstor.org/stable/j.ctt24h8h5.

human dignity, the principle of self-determination With the introduction of the European Regulation n. 2016/679 (GDPR), the EU has moved in this direction. With this in mind, the European legislator has introduced new principles, including accountability, the principle of transparency (with regard to information of interested party), adequacy, lawfulness, and proportionality. The latter requires a balance between the right to privacy and the legitimate interest of the Data Controller (e.g., national security) in order to avoid discrimination and reduce the risk of profiling.

5. Profiling Risk and actions useful to mitigate it

Profiling[32] is the most critical and insidious form of automated processing.

Due to the scope of automated decision-making[33] pursuant to article 22 GDPR, it is "a procedure which may involve a series of statistical deductions. It is often used to make predictions about people, using data from various to infer something about an individual, based on the qualities of others who appear statistically similar"[34].

It helps organizations in automation of such outcomes and results by providing large and diversified data sets (personal data like contacts, social, location) sourced from varied data sources (e.g., web portals, company web site, social media). Automated data profiling also provides more comprehensive insights for better decision making. Results of customer age and product usage profiling can be used for customized service offering and digital marketing.

It still happens that many companies often begin the profiling process without notifying or obtaining due consent form individuals from whom the data has been collected, in breach of GDPR. The same Regulation, in fact, provides that the data subject should have the right not to be subject to a decision, which may include a measure, evaluating personal aspects relating to him or her which is based solely on automated processing, and which produces legal effects concerning him or her[35]. This means that the individual can deny consent to any automated processing, including profiling, where it is deemed to have legal effects affecting him/her or significantly affecting his/her person.

32 On the definition of profiling, see article 4, par.1., n. 4, GDPR.

33 Recital no. 71 of the GDPR cites, as examples of automated decisions that can significantly affect the rights and freedoms of individuals, the automatic refusal of an online credit application or electronic hiring practices without human intervention.

34 This definition comes from Working Party article 29, Guidelines on Automated individual decision making and Profiling for the purpose of Regulation 2016/679, 17/EN, WP 251 rev.01, 3 October 2017.

35 See article 22, par.1 and 2, and Recital no. 71 GDPR, which sets out a general prohibition on solely automated individual decision-making with legal or similarly significant effects. This means that the controller should not undertake the processing described in Article 22 (1) unless one of the Article 22 (2) exceptions.

Each of us, after being informed in a clear and precise way by the Data Controller, must be able to freely decide whether to give consent to the processing of data. Therefore, if the data subject is obliged to give consent in order to use a service, this represents a violation of the GDPR. It must therefore be considered that the irresponsible and illegal use of data, also resulting from i) the absence of controls and procedures ii) the use of personal data in breach of the Data protection legislation, can lead to significant health, financial and reputational risks for individuals. In addition, individuals exposed to such risks may be discriminated against or subjected to abuse and stereotyping. In other words, Profiling could cause ethical issues. To this extent, what kind of actions are useful for private and public players to protect Privacy of the individuals and mitigate risks? GDPR, to avoid prejudice to the legal and private sphere of the individual, identifies a series of requirements to make automated processing compliant with the legislation and mitigate privacy risks. Being guided by the principles of transparency, proportionality, privacy by design[36] and privacy by default, each Data Controller, with a risk - based approach, shall, prior to the processing, carry out an assessment of the impact of the envisaged processing operations on the protection of personal data where a type of processing is likely to result in a high risk to the rights and freedoms of natural persons.

Data Protection Impact Assessment (DPIA) is an important instrument in terms of accountability as it assists the controller in compliance with the requirements of the GDPR and allows to identify the risks and appropriate measures to mitigate them.

Among the types of processing to be subjected to the DPIA, identified according to the criteria drawn up by the Working Party 29[37], in fact, there is also the one carried out using biometric facial recognition, an innovative technology that can also be used for "a systematic monitoring of a publicly accessible area"[38] or for profiling. It's therefore clear that by reflecting on the purposes of the processing and identifying risks in terms of probability and severity, countermeasures can be more easily developed to reduce them.

---

36 Privacy by design implies that data protection is integrated into the design of service or process technology from its inception. This involves *assessing and predicting the specific risks* to the personal data that will be processed. On the above-mentioned principle, IT World Canada reports that *ISO 31700* on privacy by design will be adopted on Feb. 8, 2023. On this notice, see the link below: https://iapp.org/news/b/iso-set-to-adopt-privacy-by-design-standard.

37 On the criteria, in particular no. 8 (Innovative use or applying new technological or organizational solutions), see Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, 17/EN, WP 248 rev.01, also mentioned in Annex 1 to Measure No. 467 of 11 October 2018, entitled "List of types of processing subject to the consistency mechanism to be subject to impact assessment".

38 On this concept, see Working Party article 29 Guidelines on Data Protection Officer 16/EN WP 243.

In this way, several security measures like encryption and pseudonymization, provided by article 32 GDPR and introduced in internal targeted procedures on the processing of personal data through fingerprint and retinal scanning, can be taken to make personal data untraceable to persons not authorized to read it but the most important measure is to provide a clear and accurate privacy notice with a legitimate legal basis, ensuring especially the exercise of the right to erasure[39] and the right to object.

Moreover, if an AI system is used for processing purposes, as in the case of use of video surveillance systems with biometric facial recognition technologies, with Article 13 of the Proposal of Regulation on Artificial Intelligence, the European legislator provides that information on the risks associated with a compliant or improper use of this technology would have been provided to the user. To this extent, will the implementation of such measures be enough to protect individuals from the "Big Brother who observes us"[40]?

6. Ideas and reflections

It will be a true challenge because Data are now the object of desire for reasons of control of the *Big Tech*, web giants who nowadays hold the true power of surveillance, today protagonists of a new form of "capitalism"[41]. The widespread use of biometric facial recognition in "ordinary" circumstances and for mere facilitating purposes, is generating a real danger for rights in terms of progressive loss of freedom, and underestimation of the consequences that may result from an unconscious use of such technologies. The guarantees already imposed by European Data Protection legislation are undoubtedly significant to protect right and freedom and include high criminal penalties in the case of discriminatory profiling based on particularly sensitive data, such as biometrics. However, in a digital world, where new AI systems are constantly emerging, it is essential to assess the interests involved on a case-by-case basis (e.g., security

39 As noted by Bellomia, V. (2019) 'Right to be forgotten and the Information Society', Milan, p. 25, "To ensure right to be forgotten means simultaneously protecting the most essential personal assets of individuals and contributing to the correctness of information. On the other hand, if the wealth of information is a valuable asset for the economy and businesses, on the contrary, an excess of information risks causing unnecessary and excessive management costs and organizational inefficiencies".

40 Orwell, G. (1950) *1984*, Mondadori.

41 Mobilio, G. (2021) *Face recognition technologies: risks to fundamental rights and regulatory challenges*, Editoriale Scientifica. The Authors refers to Zuboff, S. (2019), *Capitalism of Surveillance. The future of humanity in the era of new powers*, Luiss University Press, Rome. Data are collected to be transformed into "predictive products", able to determine what we will hear, think and do, to influence human behavior and transform the data into a source of gain, to be given in what is defined as the "future behavioural market", including online advertising, insurance, banking, finance, retail, etc. (18).

and privacy), possibly by balancing them, in accordance with the principles of necessity and proportionality. In order to prevent the ideal building designed by Bentham in 1791[42] from becoming the main model of social control. Adopting code of conduct, and other flexible instruments, like soft law[43], introducing a *moratorium* on the use of facial recognition in public places for a period of three to five years could be a decisive choice as it would help to verify the real impact of such technologies, as well as "to consolidate that particular identity profile that Europe is progressively affirming on the ground of the relationship between law and technology, trying to reshape it in an anthropocentric key"[44].

The European Union is moving in the right direction to affirm the primary importance of rights. The recent Proposal of Regulation on Artificial Intelligence, in fact, represents an attempt to give players guidelines to follow in the offer of AI systems, which require compliance with GDPR.

However, when in everyday life only the fingerprint or the retina scan will be needed, are we going to be ready to "not make us spy too much"[45], not to fall into the trap of "commercial surveillance"? Or will we unknowingly lose control of our data, and our privacy will be continuously compromised?

42 Panopticon, the ideal building designed by Jeremy Bentham in 1791 and subject of debate between philosophers and jurists, as Foucault, who in his most famous book *Surveiller et punir: Naissance de la prison* (1975) focuses on the historical analysis of punitive, sanction and penitentiary systems including Panoptism, a form of power that is exercised through the meticulous control of spaces.

43 On the meaning of soft law, Mostacci, E. (2008) *The soft law in the system of the sources: a comparative study*, Wolters Kluwer, cit., 41 ss. and 117 ss. Soft law could be a crossroads between the legal modalities of the rule of law and the governance modalities of globalization, which seems capable of circumventing the sovereignty of the State and its most typical manifestations of legislative will without calling them into question; see Maestri, E. (2015) 'Lex Informatica'*, Diritto Scienza Tecnologia Law Science Technology*, cit. 102. Some example of acts of soft law on biometric facial recognition: Cnil, Facial recognition: for a debate living up to the challenges, 15 November 2019; Information Commissioner's Office, ICO investigation into how the police use facial recognition technology in public places, 31 October 2019.

44 Resta, F., Pollicino, O. (2020) 'Face recognition and data protection beware of the point of no return', *Diritti Comparati*.

45 From the movie *Rear Window*, A. Hitchcock, 1954.

INTERNET LAW AND DIGITAL SOCIETY

An International Overview

Edited by Paulina Kowalicka

The first International Conference of the Information Society Law Center, held in 2023, delved into key issues at the intersection of law and new technologies, analyzing regulatory, ethical, and operational challenges in the contemporary digital context. The debate addressed four fundamental themes, attempting to provide a comprehensive view of the main legal areas involved in the technological landscape: international cybersecurity law, data protection, digital platform regulation, and the development of new technologies based on artificial intelligence systems.

Digitalization has radically transformed the way people interact, work, and learn. This revolution offers unprecedented potential to stimulate innovation, increase efficiency, and improve quality of life, while at the same time posing significant challenges that require a coherent and appropriate policy response.

The main goal of this dialogue between scholars of all over the world is to foster a better understanding of the global and local dynamics governing the digital world and the implications for law and society.

ISBN 979-12-5510-207-6 (print)
ISBN 979-12-5510-210-6 (PDF)
ISBN 979-12-5510-212-0 (EPUB)
DOI 10.54103/infolawsoc.207