ACTA ELECTRONICA UNIVERSITATIS LAPPONIENSIS 368

Emmanuel Salami

Artificial Intelligence: The end of Legal Protection of Personal Data and Intellectual Property?

Research on the countering effects of data protection and IPR on the regulation of Artificial Intelligence systems

LAPIN YLIOPISTO
UNIVERSITY OF LAPLAND

Acta electronica Universitatis Lapponiensis 368

EMMANUEL SALAMI

Artificial Intelligence: The end of Legal Protection of Personal Data and Intellectual Property?

Research on the countering effects of data protection and IPR on the regulation of Artificial Intelligence systems

Academic dissertation
to be publicly defended with the permission
of the Faculty of Law at the University of Lapland
in lecture room C 147 (former LS 1, Fellman Hall) on 12 January 2024 at 12 noon.

LAPIN YLIOPISTO
UNIVERSITY OF LAPLAND

Rovaniemi 2024

The University of Lapland
Faculty of Law

Supervised by:

Professor Rosa Maria Ballardini, University of Lapland
Emeritus Professor Rob Van den Hoven Van Genderen, Vrije Universiteit Amsterdam

Reviewed by:

Professor Burkhard Schafer, University of Edinburgh
Professor Anette Alen-Savikko, University of Helsinki

Opponent:
Professor Burkhard Schafer, University of Edinburgh

CC University of Lapland, Emmanuel Salami

Layout: Minna Komppa, Taittotalo PrintOne

Acta electronica Universitatis Lapponiensis, 368

ISBN - 978-952-337-399-0
ISSN 1796-6310

The permanent address of the publication:
http://urn.fi/URN:ISBN:978-952-337-399-0

Abstract

EMMANUEL SALAMI

Artificial Intelligence: The end of Legal Protection of Personal Data and Intellectual Property?

Research on the countering effects of data protection and IPR on the regulation of Artificial Intelligence systems

Rovaniemi, University of Lapland, 2024, 122 pages
Acta electronica Universitatis Lapponiensis, 368
ISBN - 978-952-337-399-0
ISSN 1796-6310

Artificial Intelligence systems have gained notoriety for changing (and having a great potential) to further change the way we live. The use of AI impacts the rights and freedoms of natural persons necessitating the revision of various laws relevant to AI. This research considers the intersection of data protection and intellectual property law as it impacts the rights and freedoms of natural persons. This research argues that data protection and intellectual property law interrelate in such a manner that the (non) regulation of one legal field might (negatively) impact the other. This research examines some of these issues, (including data reidentification) and further proposes the redefinition of the concept of personal data as a means of ensuring that the application of data protection and intellectual property law to AI does not limit the development, adoption, and use of AI.

Foreword and Acknowledgements

As a young student growing up in the bustling city of Lagos, Nigeria, I always imagined studying for a PhD at some point in my career. However, I never thought I would end up at the northernmost university in the European Union! My first visit to Rovaniemi was as an Erasmus student and, like many other "tourists", so allured was I that I decided to stay back. In my own case, enrolling in the doctoral program was my way to never leave Lapland. My experience at the University of Lapland has been nothing short of incredible.

Completing this thesis would never have been possible without the most amazing and empathic supervisors' support. My first supervisor, Professor Rosa Maria Ballardini, posed an inspirational figure during my master's and doctoral study – from the moment I enrolled at Lapland till date! Thank you for all you have done for me and the university community. We are indeed lucky to have you! I am also indebted to em. Professor Rob van den Hoven van Gendereen – my second supervisor – whose recommendations, nuanced guidance and critical thinking were crucial in shaping my research. I am also grateful to em. Professor Rauno Korhonen who always provided an endless supply of books that was critical to the development of the first version of my research proposal.

My profound gratitude goes to my pre-examiners, Professor Burkhard Schaffer and Professor Anette Katariina Alén, for their invaluable review and feedback, which shaped my research output. I am also grateful to Professor Schaffer for his role as an external reviewer during my halfway doctoral defence.

I want to thank all the anonymous peer reviewers of the articles that make up my research. Their new perspective helped me unravel new issues in my research. As part of my research methodology, I also had the privilege of interviewing four anonymous AI practitioners who eagerly helped me with their invaluable insights and experience. I want to express my sincere and unreserved gratitude to them for their significant contribution to my research. My research was also reviewed by Professor Alex Makulilo, to whom I am highly grateful for his time and helpful feedback.

I am also grateful to my family for their support throughout my study and stay abroad. Feeling loved from home is undoubtedly one of the foundational forms of love. I specially dedicate my work to the memory of my late father **– Makinwaye -** thank you for the solid foundation you laid!

From my experience, it takes a village to complete a PhD. I want to thank the myriad of people who have contributed in one way or another to my research,

including friends, university staff, employers, work colleagues and even project funders who gave me the wings to fly and the opportunity to broaden my horizon. I thank you all for the kind support and look forward to giving back to this fantastic community that has contributed immensely to building me up.

Rovaniemi, November 2023

Emmanuel Salami

Contents

Original Articles

The thesis is based on the following original articles, which will be referred to in the text similarly to other publications attributable to authors.

I. AI-Generated Works and Copyright Law: Towards A Union of Strange Bedfellows - Journal of Intellectual Property Law & Practice, Volume 16, Issue 2, (2021) 124–135.

II. Balancing Competing Interests in the Reidentification of AI-Generated Data - European Data Protection Law Review Vol 8 Issue 3 (2022) 362 – 376.

III. Autonomous Transport Vehicles vs The Principles of Data Protection Law: Is Compatibility Really an Impossibility? International Data Privacy Law, Volume 10, Issue 4, (2020) 330–345.

IV. AI, Big Data and The Protection of Personal Data in Medical Practice. European Pharmaceutical Law Review, Vol 3 Issue 4 (2019) 165 – 175.

All articles are reproduced with the permission of their copyright holders.

Abbreviations

A29WP - Article 29 Working Party
AI - Artificial Intelligence
AI HLEG – AI High-Level Expert Group
AC – Autonomous Cars
AS – Autonomous Ships
AV – Autonomous Vehicles
CDPA - Copyright, Designs and Patents Act
CJEU - Court of Justice of the European Union
CNIL - CommissionNationale de l'informatique et des Libertés
CRISPR-Cas9 - Clustered Regularly Interspaced Short Palindromic Repeats
DPO - Data Protection Officer
EC - European Commission
ECHR - European Convention on Human Rights
EDPB - European Data Protection Board
EDPS - European Data Protection Supervisor
EEA - European Economic Area
EU - European Union
GDPR General Data Protection Regulation
ICO - Information Commissioner's Office
IOT - Internet of Things
IP, IPR - Intellectual Property (Right)
ISP - Internet Service Provider
MVC - Mobile Vehicular Cloud
SAE - Society of Automotive Engineers
SCC - Shore Control Centre
TRIPS - The Agreement on Trade-Related Aspects of Intellectual Property Rights
WIPO – World Intellectual Property Organisation

1 Introduction to the research

Technology is governed by rules similar to those applicable in non-tech areas of human endeavour. Undoubtedly, Artificial Intelligence (AI) is one of the leading technologies of the past decade and will remain so for the foreseeable future. However, owing to their uniqueness, AI systems are challenging the fundamentals of various fields of law by raising many questions about the effectiveness of these fields of law in regulating it appropriately. This research focuses on the impact of AI systems on the areas of data protection and Intellectual Property Rights (IP, IPR) law, the rationale being the inter-relationship between both fields, which might result in regulatory discrepancies in the context of AI if one area is not concurrently regulated alongside the other. The inter-relationship between data protection and IP law stems from the ascription of IPR to data can result in unanticipated data protection concerns. This inter-relationship is explored in the articles making up this research and throughout this synthesis.

Data is an essential component in the operation of AI systems, with its relevance cutting across data protection and IP law. In fact, in using AI systems, data could be either personal or non-personal. Personal data is any data that, alone or in combination with other categories of data, leads to the identification of natural persons,[1] while non-personal data does not.[2] The definition of personal data under the GDPR is expansive and (as will be discussed subsequently) raises various challenges for AI because almost every non-personal dataset can potentially end up as personal data. Furthermore, data is of great significance to AI systems because it is used in training AI to perform its tasks (machine learning), data about (both natural and non-natural) persons are also processed and analysed by AI, and the valuable information generated from these processing activities is in itself in the form of data.

This is an article-based doctoral thesis (subsequently referred to as research) that, *among other things,* highlights some of the main concerns that can potentially arise from the use of selected AI systems within the context of data protection and IP law. The research uses the tool of four peer-reviewed articles to paint a picture of the inter-relationship between data protection and IP law. Articles I to IV cover

1 Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, GDPR) (OJ L 119/2016), Art. 4(1).

2 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, Art. 3(1).

various IP and data protection law issues but wholistically tell a story about the effects of data reidentification on both fields of law, culminating in consideration of the effectiveness of the current definition of personal data.[3] While the research answers various research questions, these research questions revolve around the inter-relationship between data protection and IP law. Therefore, the articles cover distinct data protection and IP law issues that jointly reflect the inter-relationship between both fields of law. These concerns are particularly feasible in the event of data (re)identification which results from the singling out of natural persons from either anonymised data or data initially believed to neither relate to nor identify natural persons. The research comprises IPR and data protection law articles. In addition, the IPR articles address the feasibility of ascribing IPR to AI-generated works.

On the other hand, the data protection law articles address the data protection law risks that might arise from such an ascription. The research questions considered in this research as well as their interconnectivity, are addressed subsequently in this synthesis. AI systems that generate copyrightable works, autonomous transport vehicles, and AI systems deployed in medicine are the use cases in this research.

The scope of this research falls within the framework of relevant European Union (EU) IP and data protection legislations, cases, and principles. An aggregation of all the issues (briefly identified above), which will be addressed in more detail, will reveal an overarching question about the appropriateness of the current definition of personal data in the age of AI. It is argued that the interpretation of personal data as any data which, alone or in combination with other types of data, might lead to the identification of natural persons[4] is too broad when considered within the context of the large volumes of data required to keep AI systems operational.[5] This is because the impact of technology and data reidentification techniques has blurred the line between personal and non-personal data. This means that data that is apparently (believed to be) non-personal might be capable of being further processed in a manner that results in the identification of natural persons. This is compounded by the fact that data which might lead to the identification of natural persons, when combined with other categories of data, may fall within the definition of personal data, thereby making the current definition of personal data too broad for the use of AI. This is because, with AI and data reidentification techniques, a lot of seemingly non-personal data categories can result in the identification of natural persons. Therefore, this necessitates reconsidering the definition of personal data

3 See section 1.1 of this research.

4 Art 4(1) GDPR.

5 For a further reading on the data requirements of AI, see - The Information Commissioner's Office, Big data, artificial intelligence, machine learning and data protection, 20170904, version 2.2, 12 <https:andandico.org.ukandmediaandfor-organisationsanddocumentsand2013559andbig-data-ai-ml-and-data-protection.pdf> accessed 14/07/2021.

if AI systems can legitimately access the much-needed data required to carry out their tasks.[6] This research advocates for a re-evaluation of data protection and IP law to create a practical solution that benefits the development of AI in the EU. This ensures that the regulation of IPR protection in the use of data in AI systems does not result in unanticipated negative consequences for data protection law and vice versa. In this synthesis, data reidentification refers to the singling out of natural persons from anonymised data. Still, it is also loosely used (for convenience) to refer to the identification of natural persons through pieces of non-personal data.

This research contributes to the above debate by addressing selected data protection and IP rights issues within the context of AI, identifying potential pitfalls, and proposing solutions to identified problems. These selected data protection and IPR issues will be cumulatively addressed in this research taking cognizance of relevant literature. The research is targeted at regulators, AI manufacturers, users/subscribers, researchers, students, and all other relevant stakeholders impacted in the use of AI.

1.1 Structure and content of the research

As indicated above, this article-based doctoral thesis comprises this introductory chapter (synthesis) and four peer-reviewed scientific articles published in scientific journals.

This synthesis is divided into six substantive chapters. The first part, titled "Introduction to the Research", presents the rationale, objectives, structure, and methodology adopted in this research. This chapter also considers the relationship between the four articles in more detail. Other issues considered in this chapter include a conceptual elaboration on the key topics such as an overview of autonomous vehicles, the use of AI in medicine, and an 'overview of AI systems generating copyrightable works'. These topics form important parts upon which the use cases in this research are modelled. The second chapter of the synthesis, titled "Background to the Research", reviews the existing literature relevant to this research as well as the state of relevant existing and proposed legislations. The third chapter, titled "Summary of the Essays", summarises the four published articles while highlighting their overall objectives and main contributions. Chapter four of this synthesis examines some potential implications of not concurrently regulating the fields of IP and data protection law while highlighting recent regulatory efforts and their impact on the research. Chapter five summarises the analysis and findings from the empirical study carried out in this research. Finally, the considerations and conclusions of the research, as well as its possible implications, shortcomings, and

6 The definition of personal data and its elements have been addressed in Article II of this research.

proposition for the redefinition of the concept of personal data, are considered in chapter six of this synthesis.

The four articles making up this research are divided between the fields of IP and data protection law. The first article, titled '**AI-Generated Works and Copyright Law: Towards A Union of Strange Bedfellows**' which is published in the Journal of Intellectual Property Law and Practice, considers some of the possible challenges that might be encountered while trying to protect AI-generated works through copyright law in the EU IP framework. This article focuses on the 'authorship' eligibility requirement of copyright law and considers the possibility of AI being recognised as an author within the EU IP law framework. This article lays a foundation for some of the (data reidentification issues) discussed in Article II and subsequent parts of this research. Furthermore, this research examines the effect of data subject requests when some form of sui generis IPR protects the AI-generated data. The premise for this examination is outlined in Article I, which addresses the possibility of ascribing some form of *sui generis* IPR to AI-generated works (and, by extension, AI-generated data). The critical consideration here will be whether data subject requests can be fulfilled if some form of sui generis IPR protects the AI-generated data. These considerations lead to the over-arching question of this research: the effectiveness of the definition of personal data within the context of AI and whether such definition should be expanded or further restricted. At the foundation of all these considerations is the capability of AI to own or hold IPR in its work, which is established in Article I.

The second article, titled **'Balancing Competing Interests in the Reidentification of AI-Generated Data',** has been published in the European Data Protection Law Review. The theme of this article revolves around the implication of the (re)identification of AI-generated data protected by some form of *sui generis* IPR. Some propositions which ensure a balance between the protection of IPR and data protection rights are also explored in this article. This article is critical to the overarching question of this research because the transformation of personal data to non-personal data and vice versa is fundamental to the examination of the effectiveness of the current definition of personal data within the context of AI. The third article is published in the International Data Protection Law Journal and is titled **'Autonomous Transport Vehicles vs the Principles of Data Protection Law: Is Compatibility Really an Impossibility?'** This article considers selected data protection concerns (which can potentially arise) in using autonomous vehicles and how these can be resolved for all stakeholders, including data controllers, pedestrians, regulatory authorities, etc. The fourth and final article, '**AI, Big Data and The Protection of Personal Data in Medical Practice**', is published in the European Pharmaceutical Law Journal. This article examines selected data protection law considerations, including concerns relating to data reidentification, which (can potentially) arise in the use of AI systems in (tele)medicine. Articles

III and IV only address possible data protection law concerns in using AI systems to typify some risks that could arise upon data reidentification. Considering these data protection concerns determines the necessity or otherwise of the overarching research question, which pertains to revisiting the definition of personal data.

1.2 The scope of AI within the context of this research

AI has been defined in several ways based on the diverse perspectives of various authors, policymakers, stakeholders, etc.[7] One definition of AI is its description as the process where machines display some level of intelligence in the performance of their tasks.[8] Russel and Norvig define AI as 'the study of agents that exist in an environment and perceive and act'.[9] On his part, Finlay defines AI as the replication of human analytical and/or decision-making capabilities.[10] The term machine learning is sometimes used interchangeably with AI, even though they mean different things. Machine learning uses algorithms to analyse data by discovering functional patterns within data sets.[11] AI systems rely on machine learning processes to learn how to perform tasks in a manner analogous to human learning. Therefore, while AI systems are designed to carry out tasks with varying levels of intelligence, machine learning is the procedure through which the performance of the tasks is learnt and improved upon by AI.

Even though AI systems are expected to display human-like intelligence, this is at best an aspiration as at the time of writing. The limitation of AI's simulation of human-like intelligence is noted through classifying AI systems into weak or narrow and strong or general AI systems. The terms weak and strong AI systems will be referenced throughout this synthesis. Strong AI is synonymous with autonomous systems which can perform (a variety of) tasks autonomously and with little supervision. In contrast, weak AI systems typically require more human intervention to accomplish defined tasks.[12] Strong AI is largely only a product of science fiction.

---

7 Stuart J. Russell and Peter Norvig, Artificial Intelligence: A Modern Approach (Pearson, 3rd edn, 2010) 7; Alan M. Turing, 'Computing Machinery and Intelligence' (1950) Mind 49 433-460. Article 3 (1) 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 <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX-%3A52021PC0206> accessed 30/07/2022.

8 Jerry Kaplan, Artificial Intelligence what everyone needs to know (Oxford publishers, 2016) 1.

9 Peter Norvig and Stuart Jonathan Russell (n 7) 7.

10 Steven Finlay, Artificial Intelligence and Machine Learning for Business: A No-Nonsense Guide to Data Driven Technologies (Relativistic, 3rd edn, 2018) 10.

11 Ibid.

12 Jonathan Flowers, 'Strong and Weak AI: Deweyan Considerations' (2019) AAAI Spring Symposium: Towards Conscious AI Systems <https://api.semanticscholar.org/CorpusID:57663042> accessed 11/07/2021.

Depending on the context, this research refers to both strong and weak AI, with distinctions subject to necessity. There are currently various ways through which this limitation of AI is defined. One such way is the classification of autonomy in cars by the Society of Automotive Engineers (SAE). According to the SAE's classification, level 1 involves total human control and excludes any automation, levels 2-4 consist of some automation with varying levels of human intervention, while level 5 involves complete automation without human intervention.[13]

Selected AI systems are used to highlight the issues considered in this research within the fields of data protection and IP law. The AI systems relied upon practically highlight relevant issues considered in this research, thereby supporting an understanding of the said issues. These AI systems are espoused as follows:

1.2.1 Autonomous Vehicles

Autonomous Vehicles are a significant type of AI with a vast potential to revolutionise various industries, particularly the automotive industry.[14] Autonomous Vehicles are referenced widely in this research, and include autonomous cars, autonomous ships, autonomous trains, and autonomous planes.[15] As earlier stated, five levels of automation requiring varying levels of human intervention have been identified for autonomous cars.[16] It is level 5 which is the fully automated and autonomous driving stage that is referenced in this research. Therefore, autonomous vehicles in this research embody the most advanced (and futuristic) stage of vehicle automation, where the vehicle is always independently responsible for all driving tasks without any human intervention.[17] The mode of data collection in autonomous vehicles is traditionally through sensors and cameras placed at strategic parts of the vehicle.[18] The use of autonomous vehicles in this research is integral to amplifying the data protection concerns that may arise from the use of AI systems generally. Some of the standard features of AI systems such as the possibility of collecting and processing

13 Society of Automotive Engineers, 'SAE Levels of Driving Automation Refined for Clarity and International Audience' (SAE website, 3 May 2021) <https://www.sae.org/blog/sae-j3016-update> accessed 18/10/2022.

14 CBInsights, '33 Industries other than Auto that Driverless Cars could turn upside down' (CBInsights, 2018) <https://www.cbinsights.com/research/13-industries-disrupted-driverless-cars/> accessed 26/07/2021.

15 The articles focus on autonomous cars and autonomous ships which are more advanced AV types than autonomous trains and autonomous planes.

16 Society of Automotive Engineers (n 13).

17 For further reading, see Andreas Herrmann, Walter Brenner and Rupert Stadler, Autonomous Driving: How the Driverless Revolution will Change the World (Emerald Group Publishing 2018) 3, 8-9, 47-51.

18 Ibid 95-96. See also Mogens Blanke, Michael Henriques, Jakob Bang, 'A pre-analysis on autonomous ships' (Technical University of Denmark, 2016) 1 < https://www.semanticscholar.org/paper/A-pre-analysis-on-autonomous-ships-Summary/4eabcca691a52956f697f560dca3c1ce942781d8> accessed 22/06/2021.

more data than initially anticipated, having multiple third-party processors, difficulty in the provision of information about the nature of the processing activity, difficulty in the determination of a legal basis, etc. are all manifest in the use of autonomous vehicles. Article III of this research specifically addresses the data protection concerns of autonomous vehicles. The consideration of autonomous vehicles in this research is relevant mainly because of their popularity, momentum, and acceptance.[19] Autonomous vehicles are also being tested globally, including within the EU.[20]

1.2.2 AI Systems in Medicine

AI also plays a massive role in the practice and provision of medical services. Medical AI devices have been developed to support medical practitioners in disease diagnosis,[21] drug discovery,[22] telerobotic surgeries, robot-assisted surgeries, etc. [23] As with autonomous vehicles, medical AI systems have been used to highlight the data protection (and, where applicable, IPR) concerns that flow from its use. The data protection-relevant features of medical AI devices, including the involvement of multiple stakeholders in its usage and the multi-jurisdictional possibilities of its usage,[24] the fact that these devices process large volumes of sensitive personal data, etc.,[25] are all considered in this research. The sensitive nature of the personal data

19 A lot of auto manufacturers and even traditional tech-based companies have transitioned into AV manufacturing. See CBInsights, '40+ Corporations Working on Autonomous Vehicles' (CBInsights, 2020) <https://www.cbinsights.com/research/autonomous-driverless-vehicles-corporations-list/> accessed 22/07/2021.

20 Autonomous vehicles have been tested in China, and some major US cities. See Abigail Ng, 'Completely driverless cars are being tested in China for the first time' (CNBC, 2020) <https://cnb.cx/37CgG9J> accessed 22/07/2021. Germany is also planning to partly roll out AV in 2022. See Rebecca Bellan, 'Germany gives greenlight to driverless vehicles on public roads' (Techcrunch, 2021) <https://techcrunch.com/2021/05/24/germany-gives-greenlight-to-driverless-vehicles-on-public-roads/> accessed 22/07/2021.

21 Ali Madani, Ramy Arnaout, Mohammad Mofrad and Rima Arnaout, 'Fast and Accurate View Classification of Echocardiograms Using Deep Learning' (Digital Science 1, 2018) 6.

22 Lothar Terfloth, Simon Spycher, Johann Gasteiger, 'Drug Discovery: An Overview' in Thomas Engel and Johann Gasteiger (eds), Applied Chemoinformatics (Wiley, 2018).

23 Bernard Dickens and Rebecca Cook, 'Legal and Ethical Issues in Telemedicine and Robotics' (2006) 94 International Journal of Gynecology & Obstetrics 1, 73-78.

24 Telerobotic surgeries are carried out by multiple parties possibly across multiple locations. Telerobotic surgeries are carried out by robotically controlled instruments through which coronary intervention can be undertaken by a doctor without any physical contact with the patient. See, Mahesh Langa, 'Ahmedabad Doctor Performs Telerobotic Surgery on Patient 32 km Away' (The Hindu, December 2018) <https://www.thehindu.com/news/national/other-states/ahmedabad-doctor-performs-telerobotic-surgery-on-patient-32-km-away/article25675166.ece> accessed 22/07/2021.

25 Sensitive personal data is data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, and the processing of genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation. See Art 9 (1) GDPR.

being processed by AI systems in the medical sector makes these AI systems a good use case for projecting the objectives of this research.

1.2.3 AI Systems Generating Copyrightable Works

AI has significantly contributed to the creative industry, particularly in creating cinematographic, artistic, literary, and musical works. These AI systems can create works which would have been eligible for copyright protection had they been created by human authors. Some AI systems which are actively used in the creation of various copyrightable works were used to highlight selected IPR issues, including but not limited to the legal possibility of AI systems owning/holding IP rights, the pros and cons of letting AI-generated works lie in the public domain; the feasibility of legal personhood as a means of granting AI systems the needed personhood to own/hold AI-generated works; etc. Some notable AI systems considered in the research include Benjamin the bot,[26] Next Rembrandt,[27] and AICAN.[28] These AI systems were selected because they are leading AI systems that create works that help bolster this research's objectives.

1.3 Aim and Research Questions

As previously highlighted, this research seeks to resolve selected data protection and intellectual property law conflicts that can potentially arise in the deployment of AI systems. The resolution of these potential conflicts leads to the overarching issue/question of this research, which is the appropriateness of the current definition of personal data in light of the data requirements of AI systems. This overarching research question (which is subsequently addressed in further detail) revolves around the suitability of the definition of personal data to either encourage or hinder the development of AI.[29] Before arriving at this overarching question of the research, some other issues and research questions form the building blocks that cumulatively help in achieving the objective of the research. These issues and research questions are reflected in the four articles, constituting

26 'Benjamin' is the movie director who was responsible for directing the movie 'zone out' by piecing thousands of old movies together. See Lauren Goode, 'AI made a movie and the results are horrifyingly amazing' (Wired, 2018) <https://www.wired.com/story/ai-filmmaker-zone-out/> accessed 01/06/2021.

27 Andres Guadamuz, 'Artificial intelligence and copyright' (WIPO magazine, October 2017) <https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html> accessed 04/09/2021.

28 AICAN creates art based on images dating back to five centuries ago with which it had been fed during machine learning. Ahmed Elgammal, Meet AICAN, a machine that operates as an autonomous artist' (The Conversation, October 2018) <https://theconversation.com/meet-aican-a-machine-that-operates-as-an-autonomous-artist-104381> accessed 17/10/2020.

29 See chapter 6 of this synthesis.

essential components of this research and laying the foundation for addressing the overarching research question.

Some of the research questions identifiable from the articles are listed as follows:

Article I - Can AI-generated works produced by AI systems be eligible for copyright protection?; What are the legal consequences of ascribing copyright protection to AI-generated works?; Should AI-generated works lie in the 'public domain' rather than ascribe copyright protection to them?; Can legal personhood be ascribed to AI systems?

Article II - Can AI-generated non-personal data be subject to IPR or *sui generis* protection?; Can natural persons be singled out from apparently non-personal data?; What are the data protection concerns that might arise in the reidentification of AI-generated data?; Who is an appropriate party to take responsibility for non-personal data that results in the (re)identification of natural persons, and how can these data protection risks be averted?

Article III - What data protection concerns can arise from using autonomous vehicles within the EU?; Based on these risks, can autonomous vehicles comply with the principles of EU data protection law?; How can the data transfer issues that can arise in the use of AI for medicine be resolved?; Can data controllers be said to have a legitimate interest in collecting data for pedestrians on the streets?

Article IV – How can the data protection rights of medical patients be protected within the context of medical AI devices? Who takes responsibility for the data protection concerns that might arise when multiple Internet of Things (IoT) suppliers are used in telemedicine?; How can the data access/transfer issues occurring in using AI for medicine be resolved?; Should developers of AI systems used in medicine be held to a higher standard due to the sensitivity of their activity?

The research questions in Article I lay the foundation for examining the possibility of AI systems owning AI-generated works. The research questions in Article II follow closely by considering the data reidentification possibilities that might arise from AI owning AI-generated works. Finally, articles III and IV generally address data protection concerns that can potentially occur from using selected AI systems. All these issues, though distinctive, are interconnected by contributing to the resolution of the overarching research question.

1.4 The Relationship Between the Four Articles

The four articles that form part of this research contribute variously to the research theme, which revolves around the inherent data protection concerns that might arise from attributing IP rights to AI-generated works/data. The articles are divided into two IPR-themed articles and two data protection-themed articles. Collectively,

the articles tell the story of how IPR can be extended to AI-generated works and AI-generated data, focusing on the challenges encountered therein.

The first article in the research, titled "AI-Generated Works and Copyright Law: Towards A Union of Strange Bedfellows", considers the possibility of ascribing copyright protection in AI-generated works to the AI systems which have generated them. This article addresses the challenges (including the authorship eligibility requirements for AI) that such an attempt might encounter. The essence of this article is to lay a foundation for the ownership rights of AI systems either through traditional IP rights or some other form of *sui generis* right. This is needed to establish the possibility of AI owning IPR, a scenario relied upon throughout this research. The second article, titled "Balancing Competing Interests in the Reidentification of AI-Generated Data", addresses the data protection concerns that might arise in the event of the ascription of IPR to AI-generated data. This article builds on the proposition that AI systems should be ascribed with some form of IPR, as considered in article I. In contributing to the narrative of this research, article II considers practical data protection consequences that might arise should some form of IPR in AI-generated data vest in AI systems.[30] These two articles make up the IP law component of this research.

The third article, "Autonomous Transport Vehicles vs The Principles of Data Protection Law: Is Compatibility really an Impossibility?" specifically addresses data protection concerns arising in using AV. This article is linked to article II because it further highlights the data protection concerns inherent in using AI. This article aids the appreciation of the risks identified in Article II, which the substantive research seeks to identify and prevent. The fourth article, titled "AI, Big Data and The Protection of Personal Data in Medical Practice", is similar and linked to article III because it also considers the data protection concerns that arise in the use of AI, in this case, AI systems that are used in the practice of medicine. However, this article is differentiated from article III because it highlights the data protection risks inherent in using AI within the context of the special categories of personal data.[31]

Therefore, the four articles tell a combined story of how IP rights may vest in AI systems to protect their AI-generated works and the data protection risks that might arise. Undoubtedly, some of the issues considered in the articles are futuristic. However, these articles help achieve the research's aim by examining possible consequences of the non-concurrent regulation or consideration of IPR and data protection law in using AI. Furthermore, these articles highlight some of the future challenges that might arise if IPR is ascribed to AI without consideration of relevant data protection consequences.

30 For instance, article II considers the concerns that might arise in the event of data access requests in respect of AI-generated data protected by IPR. See section 3 of this synthesis.

31 Art 4 and 9 GDPR.

1.5 Research Methodology

Research methodology requires defining a problem, forming an appropriate thesis statement, collecting data, analysing said data, and making conclusions to achieve the research objectives.[32] In simple terms, research methodology identifies how research questions are resolved. Therefore, considerations such as the types of data to be collected, considered and/or ignored; how the data is to be collected; from whom it is to be collected, and the approach to its processing and analysis is determined subject to the research methodology that is to be adopted. To this end, this research utilizes a combination of the doctrinal research method, legal empiricism, and the legal design research method to analyse and convey the findings from the research questions considered.[33]

1.5.1 The Doctrinal Research Method

The doctrinal research method (also referred to as the 'black letter' methodology) focuses on the analysis of the law in legal sources such as legislations and cases.[34] This research method is known to be particularly useful when the research involves areas of law that are the subject matter of legislative instruments. Effective use of this research method requires a critical analysis of relevant laws to support or disprove the thesis statement of a given research. In addition, the doctrinal research method is used to discover the law's position on a given research topic.[35] Legal sources to be analysed in the use of the doctrinal research method include the letters of the law as stipulated in a legal instrument, judicial interpretations of the law, scholarly literature commenting on the laws, and any other relevant document capable of providing some sort of context to what the law is.

Having established the importance of legal analysis to the doctrinal research method, it is necessary to consider the deductive and inductive reasoning techniques, which are both critical, logical reasoning techniques deployed in carrying out the

32 Derek Jansen and Kerryn Warren, 'What (Exactly) Is Research Methodology? A Plain-Language Explanation & Definition (With Examples)' (Gradcoach, June 2020) <https://gradcoach.com/what-is-research-methodology/> accessed 06/08/2022. For further reading, see Reva Berman Brown, 'Doing your dissertation in business and management' (SAGE Publications, 2006) <https://dx.doi.org/10.4135/9781849209069> accessed 06/08/2022

33 Law and technology has been studied jointly under fields of law such as legal informatics, legal technology, etc. For further reading, see Sandra Erdelez and Sheila O'Hare, 'Legal informatics: application of information technology in law' (1997) 32 Annual Review of Information Science and Technology 367-402. For further readings on the history and origin of legal informatics, see Peter Seipel, Computing Law. Perspectives on a New Legal Discipline, (Liberfolag, 1977).

34 Research Guides: Legal Dissertation: Research and Writing Guide (libguides.com). <https://law.indiana.libguides.com/dissertationguide#:~:text=Doctrinal,%2C%20statutes%2C%20or%20regulations> accessed 08/08/2022.

35 Paul Chynoweth, 'Legal Research' in Andrew Knight and Les Ruddock (eds), Advanced Research Methods in the Built Environment (Wiley-Blackwell, 2008) 30.

legal analysis required in the use of the doctrinal method. Deductive reasoning is usually deployed when the legal principle being analysed can be gleaned from the statute. In such a scenario, the reasoning from one or more legal rules will lead to a logical conclusion.[36] Inductive reasoning is typically used in cases where a legal principle applies to selected use cases, thereby requiring analogy to determine if there are other cases where the legal principle has been involved similarly to that considered in the relevant research.[37]

This synthesis analyses relevant primary and secondary legal sources in data protection and IP law to resolve relevant research questions where possible.[38] The four peer-reviewed articles of this research actively deploy the tool of legal analysis that the doctrinal method offers. For instance, in considering the possibility of AI being recognised as an author in Article I, the doctrinal research method is used to evaluate the traditional criteria for authorship (in human-authored works) by analysing these criteria through the applicable EU and national copyright laws, decisions of the Court of Justice of the European Union (CJEU), national courts and scholarly literature. The cardinal issue in Article II, which revolves around the reidentification of non-personal data, is also addressed through EU and national copyright laws, decisions of the CJEU, national courts and scholarly texts. In that article, the definition of personal data was crucial for determining the possibility of data reidentification. The doctrinal method was used to analyse the definition of personal data within the scope of the GDPR, decisions of the CJEU, opinions of data protection supervisory authorities and academic articles.

This approach is also used in Articles III and IV when highlighting the data protection concerns potentially inherent in selected AI systems.

In some cases, the non-EU legal sources of law are also analysed as part of this research. An example of this can be found in Article II, which considers US court decisions in outlining cases where (non-personal) data has been ascribed with some form of IPR protection.[39] The doctrinal approach is fundamental to this research because of the abundance of statutes, cases and legal literature about the subject matter of the research. Therefore, this research will be incomplete without analysing these existing statutory, judicial, and scholarly materials. The legal documents analysed in different parts of this research have been selected to ensure that existing

36 Ibid, 32-34. For further reading, see Terry Hutchinson, 'The Doctrinal Method: Incorporating Interdisciplinary Methods in Reforming the Law' (2015) Erasmus Law Review, 8, 3. <https://ssrn.com/abstract=2734131> accessed 08/08/2022. See also Suzanne Egan, 'The Doctrinal Approach in International Human Rights Scholarship' (UCD Working Papers in Law, 29 November 2017) Criminology & Socio-Legal Studies Research Paper No. 19/17.

37 Paul Chynoweth (n 35) 32-34.

38 Terry Hutchinson, 'Valé Bunny Watson? Law Librarians, Law Libraries and Legal Research in the Post-Internet Era' (2014) Law Library Journal (106(4) 584.

39 Lowry v Legg Mason, 271 F Supp 2d 737 (D Md 2003).

schools of thought on the relevant subject matter are not omitted. The diverse nature of the data protection and intellectual property law topics involved herein makes the doctrinal research method a very appropriate and effective tool for approaching this research. The doctrinal method will also be used in this synthesis, particularly as it pertains to the overarching research question, which considers the suitability of the current definition of personal data in the age of AI.

In relying on the doctrinal research method, (the articles in), this research leans on both the deductive and inductive reasoning techniques. Deductive reasoning has been used to reach a logically inevitable conclusion from one or more legal rules. For instance, this approach is adopted in Article I, where a case is made for authorship rights to be attributed to AI-generated works even though AI lacks one of the main eligibility requirements of authorship – the requirement that an author should be a natural and/or a legal person.[40] In adopting the deductive approach, the article analysed relevant EU laws and court decisions on the definition of authorship. It made a case for attributing authorship rights to AI based on the appropriate legal sources. This research also relies on inductive reasoning to determine whether a particular legal principle applies to a set of facts or hypothetical scenarios. This reasoning technique is adopted in this synthesis where a hypothetical scenario involving data subject requests[41] made concerning IPR-protected AI-generated data is considered.[42] In Article II, the inductive reasoning technique is also used to assess the existing attitude of the courts to balancing data reidentification concerns as they pertain to data protection and IP law. In both scenarios, CJEU decisions which upheld a fair balancing of the rights to privacy and intellectual property rights, were used to preempt the applicability of existing laws and principles to the said scenarios.[43]

The limitation of the application of the doctrinal research method in this research lies in the fact that some of the issues addressed in this research are yet to become part of a recognised body of laws, even though there might be secondary sources of law in such matters. An example of such an issue is the IPR protection of AI-generated works, which is yet to be legislated upon but has received the attention of stakeholders, particularly IP regulators and scholars. In such cases, in addition to a reliance on scholarly literature, other research methods such as interviews conducted

40 For further readings on the copyright eligibility requirements, see: Rosa Ballardini, Kan He, and Teemu Roos, 'AI_Generated Content: Authorship and Inventorship in the age of Artificial Intelligence' <https://www.cs.helsinki.fi/u/ttonteri/pub/aicontent2018.pdf> accessed 04/10/2021.
Lionel Bently and Brad Sherman, Intellectual Property Law (4th Edn, OUP 2014) 93-108.

41 Chapter 3 GDPR.

42 Chapter 5 of this synthesis.

43 C-275/06, Productores de Música de España (Promusicae) v. Telefónica de España SAU [GC], 29 January 2008, paras. 62 – 68; C-461/10, Bonnier Audio AB, Earbooks AB, Norstedts Förlagsgrupp AB, Piratförlaget AB, Storyside AB v. Perfect Communication Sweden AB, 19 April 2012.

in the use of legal empiricism (and which are considered subsequently), and legal design methodology are relied upon to adequately fill the vacuum occasioned by the absence of binding legal provisions. The doctrinal research method has been criticized for encouraging a narrow interpretation of the law because of the focus on limited legal sources, possibly shutting out other legal questions that might arise from considering other research tools.[44] This synthesis tries to address this criticism by using additional research methods to balance out the doctrinal research method and obtain further perspective.

1.5.2 Legal Empiricism

Legal empiricism involves the research of legal principles and processes using social research methods, such as interviews, observations, or questionnaires. Put differently, legal empiricism consists of the use of those research methods more commonly used in the social sciences, particularly in disciplines such as economics, political science, psychology, sociology, etc.[45] This research method involves testing a theory using techniques developed in the social sciences. [46] It enhances the appreciation of theories which might justify relevant propositions.[47] Therefore, interviews have explicitly been conducted to test selected hypotheses and proposals within the scope of this research. Legal empiricism is exploratory because it permits the sampling of varying perspectives, an approach particularly useful in parts of this research where the law remains uncertain. This research method has been adopted to identify possible industry solutions and practices adopted by industry practitioners and participants in the absence of substantive legislation. To this end, this research uses explicit interviews to seek expert and industry perspectives on relevant topics.

The interviews were semi-structured, and the same questions were posed to all the interviewees. The same interview protocol was sent to all interviewees before their respective interviews.[48] Interviewees were provided with all necessary information about the research, their consent was obtained to interview via email, and all other relevant ethical considerations were considered and complied with throughout the

44 Pauline Westerman, 'Open or Autonomous: The Debate on Legal Methodology as a Reflection of the Debate on Law' in Mark Van Hoecke (ed), Methodologies of Legal Research: Which Kind of Method for What Kind of Discipline? (Hart Publishing Ltd, 2011) 86.

45 D. James Greiner, The new legal empiricism and its application to access-to-justice inquiries. Daedalus 2019, 148 (1): 64–74.

46 Mandy Burton, 'Doing Empirical research, Exploring the decision making of Magistrates and Juries' in Dawn Watkins and Mandy Burton (eds), Research Methods in Law (Routledge, 2nd edn, 2018) 55. See also Tracey George, 'An Empirical Study of Empirical Legal Scholarship: The Top Law Schools' (2005) Vanderbilt Law and Economics Research Paper No. 05-20 <https://ssrn.com/abstract=775864> accessed 04/02/2022.

47 Ibid. See also Nigel Gilbert, 'Research, Theory and Method' in Nigel Gilbert (ed), Researching Social Life (Sage, 3rd edn, 2008).

48 Appendix 1 of this synthesis.

interview and research process.[49] The interview questions were drafted to seek the perspective and current practices of industry practitioners on issues addressed in (the four published articles of) this research. The interviewees were allowed to provide answers to questions within the framework of their daily work and across their extensive knowledge of the relevant field(s). Though clearly defined, the interview questions were also open-ended, encouraging interviewees to answer questions freely without restrictions.[50] Interviewees were selected based on their experience in the relevant matters discussed in this research. The anonymity of interviewees and their affiliations has been maintained as such disclosures are not needed to achieve this research method's purpose. In addition, this research complies with the principles of ethical review and responsible conduct of research at the University of Lapland.[51]

As deducible from the interview protocol, the interviews were divided into three parts which covered the fields of data protection and intellectual property law as follows - the general questions (part A), the right to data portability (part B), and intellectual property-related questions (part C). Part A of the interview protocol provides insight into the type of AI system and data with which the interviewee has worked. At the same time, parts B and C, respectively, pose varying data protection and IPR questions, which are relevant to this research. The questions in the interview protocol can be traced to issues that have been (or will be) variously addressed in the four published articles and this synthesis. For instance, under part B of the interview protocol, the questions seek to learn from the interviewees' experience as it pertains to possible data protection issues arising in the exercise of data subject rights (particularly data portability) when IPR protects relevant (AI-generated) data. This interview question is traceable to Article II of this research, which considers, among other things, the possibility of data reidentification in the use of IPR-protected AI-generated data, especially when data subjects try to exercise their data subject rights. Furthermore, part C of the interview protocol poses questions relating to the ascription of IPR to AI-generated data to determine the attitude of companies to appropriating IP rights of AI-generated works, thereby testing relevant contentions and propositions that have been advanced in Article I of this research.

The interviews are used to evaluate the theories which underlie the various propositions advanced in this research in the light of the expert opinion of the interviewees. This is achieved by considering these proposed theories to justify their rationality, after which their outcome is then examined within the context of the

49 Martin Bulmer, 'The Ethics of Social Research' in Nigel Gilbert (ed), Researching Social Life, (London: Sage, 3rd edn, 2008).

50 For further readings on interview structures, see Jennifer Mason, 'Linking qualitative and quantitative data analysis' in Alan Bryman, Robert G. Burgess (eds), Analysing qualitative data (Routledge, 1994) 89–110.

51 <https://www.ulapland.fi/EN/Research/Responsible-research/Responsible-conduct-of-research> Accessed 16/01/2023.

views of industry participants to arrive at the best possible outcome. This ensures that the theories underlying some of the propositions in this research are balanced and considered within the context of industry realities.[52] Finally, an analysis of the interviews is carried out in subsequent parts of this synthesis.[53]

1.5.3 Legal Design Research Method

As a research methodology, legal design evolved from the design thinking methodology after some scholars adapted the design thinking methodology to the field of law.[54] Design thinking is generally defined as an analytic and creative process that engages a person in opportunities to experiment, create and prototype models, gather feedback, and redesign.[55] The goal of design thinking is the provision of a solution-based approach to problem-solving. Legal design has been defined as applying human-centred design to the world of law to make legal systems and services more human-centric, usable, and satisfying.[56] Legal design focuses on creating legal services, focusing on how functional, useful, and engaging these services are.[57] The crux of this research methodology is to apply design to the law to ensure that legal issues are resolved subject to the needs of the end users. Legal design provides better communication, customer-centric product designs, improved product offerings, etc. This research uses infographics and other media to visually represent information, such as charts, diagrams, tables, etc., to convey its findings. This research method is beneficial to this research because of its multidisciplinary nature, which cumulatively addresses issues in both data protection and IP law within the field of AI. Legal design is therefore used to effectively communicate the theme of this research to its divergent audience, which includes regulators (data protection and IP), lawyers, technologists, students, etc.

This research method has been used in this research for two primary purposes, which are simplified communication and effective end-user-focused solutioning. Simplified communication is critical to this research because of its multidisciplinary nature. This requires the discussion of legal concepts and issues in data protection

52 In using the empirical methodology generally, researchers are advised to keep their hypothesis flexible and amenable to realities that are discovered during their empirical analysis. See Mandy Burton (n 46) 56. Martyn Hammersley and Paul Atkinson, Ethnography: Principles in Practice (Routledge, 3rd edn, 1995).

53 Chapter 5 of this synthesis.

54 Apolline Le Gall, 'Legal Design beyond design thinking: processes and effects of the four spaces of design practices for the legal field' in Rossana Ducato and Alain Strowel (eds), Legal Design Perspectives: Theoretical and Practical Insights from the Field (Ledizioni, 1st edn, 2021) 29.

55 For further readings on design thinking, see Rim Razzouk and Valerie Shute, 'What Is Design Thinking and Why Is It Important?' (2012) 82(3) Review of Educational Research 330–348.

56 Margaret Hagan, 'Law by design' (2016) <https://lawbydesign.co/legal-design/> accessed 28/8/2022.

57 Ibid.

and IP law, thereby necessitating simplified language to enhance comprehension for both legal and non-legal experts alike. In its use of the legal design research method, this research uses simplified texts and communication techniques to simplify and effectively communicate its theme. For instance, Article I deploys a table to itemize different types of AI systems, their creative works and the classification of such works in an IP law context.

On the other hand, Article II uses charts to effectively indicate the transformation of data from personal to non-personal data and back to personal data. In Article III, the data collection procedures and the connected IOT in autonomous cars and autonomous ships are depicted using diagrams, enhancing a simplified comprehension requiring no technical background. In respect of the second purpose of using this research method (i.e. end user-focused solutioning), this research considers end-user interests in proposing possible solutions. This is visible in the propositions for how the transparency principle of data protection law can be implemented and complied with in AI systems.[58] For instance, in Article III, a diagram was designed to effectively provide a transparent communication of information to data subjects following the requirements of Articles 13 and 14 of the GDPR. Based on the above, this research method helps to effectively achieve the goals of this research by ensuring that the research findings are understandable and valuable to its audience. Furthermore, this research method complements the doctrinal research method, especially concerning the proposition of solutions in situations where laws are yet to be adopted.

58 Art 5 (1) (a) GDPR, Art 5 (4)(a) Modernised Convention 108.

2 Background of the research

AI-driven technologies have transformed the economy because of the increased volumes of personal and non-personal data they process. While the GDPR strictly regulates the processing of personal data by both AI and non-AI actors, the same cannot be said of non-personal data, which, at the time of writing this synthesis, is the subject matter of some proposed EU legislations.[59] Using AI-driven technologies coupled with various IOT means that large volumes of data are shared and exchanged among consumers, businesses, and society. This data flow drives innovation and competitiveness among stakeholder institutions (such as business entities and government),[60] thereby necessitating its protection from unauthorised third-party access. Expectedly, these voluminous (non-personal) data-generating AI-driven technologies raise exciting questions that form the theme of this research.

This research examines both data protection law and IP law concurrently because of the interrelationship between both fields of law, which could result in overlapping consequences for data. The research focuses on the intersection between data protection and IP law stemming from the possible ascription of IPR to (non)-personal data, which might raise some data protection concerns. The crux of this intersection lies in the fact that any attempt to ascribe IPR to AI-generated creations (including AI-generated data) might result in data protection concerns in the event of the (re)identification of (apparently) non-personal data. Such data reidentification will result in the applicability of data protection law in unanticipated circumstances. The relationship between data protection and IP law in this research stems from the possibility and consequence of such data reidentification. This might require data protection and IP law to be regulated concurrently to prevent such an outcome. Otherwise, an innocuous attempt to protect AI-generated data might raise unanticipated data protection law concerns. The possibility of such an unprecedented outcome will be considered in this synthesis using recent selected EU proposals seeking to regulate non-personal data.[61]

One significance of addressing data protection and IPR in this research is the holistic resolution of complementary issues. This approach is pre-emptive rather

59 Proposal for a Regulation of the European Parliament and of the Council on European data governance (Data Governance Act), (Proposal for the Data Governance Act, Proposal) (COM/2020/767); Proposal for a a Regulation of the European Parliament and of the Council on harmonised rules on fair access to and use of data (Data Act), (Proposal for the Data Act, Proposal) (COM/2022/68).

60 The Proposal for the Data Governance Act seeks *inter alia* to provide data for public use.

61 Footnote 59.

than reactive and helps anticipate any possible consequences that might result from regulating data protection and/or IPR. By so doing, the joint consideration of both legal fields reduces any surprises in the form of overflowing risks. Furthermore, this approach of jointly considering the areas of data protection and IP law in this research is advantageous because it reduces the possibility of overlooking complimentary issues in data regulation. One further consideration which underscores the relationship between the fields of data protection law and IPR in this research is that both legal fields are considered from the perspective of protecting IPR without jeopardizing personal data protection. In other words, as far as this research is concerned, the subject matter being considered in both fields of law is essentially data.

2.1 Existing research

As indicated above, the multidisciplinary nature of this research means that legal sources in data protection law and IP law must be considered. While the law (might) remain(s) uncertain, there is no shortage of scholarly literature on some of the issues relevant to this research. One of the critical components of this research is the consideration of the possible ascription of IPR protection (particularly copyright) to AI-generated works, as considered in Article I. The critical research question which Article I attempted to resolve was the proposition of appropriate techniques to incentivise AI-generated works. Though the law remains largely unsettled, some scholarly publications (considered in this research) propose possible solutions. Some of these solutions were proposed in the article of Aplin and Pasqualetto, stating, *among other things,* that AI-generated works should lie in the public domain because copyright protection does not necessarily incentivise authors and creatives.[62] Samuelson's position that copyright should be vested in the user of AI systems is another relevant publication which informed some of the analysis in Article I of this research.[63] These articles and others represented some existing research on the copyright protection of AI-generated works by AI systems.[64] They were variously considered in Article I of this research. The main proposition for resolving the key research question in Article I of this research was the ascription of AI systems with legal personhood. In making this proposition,

62 Tanya Aplin and Giulia Pasqualetto, 'Artificial Intelligence and Copyright Protection' in Rosa Ballardini, Petri Kuoppamäki, Olli Pitkänen (eds), Regulating Industrial Internet Through IPR, Data Protection and Competition Law (Kluwer, 2019).

63 Pamela Samuelson, 'AI Authorship?' (Communications of the ACM, 2020) Vol 63, Number 7, 22. <https://cacm.acm.org/magazines/2020/7/245693-ai-authorship/abstract> accessed 18/06/2021.

64 Lawrence B. Solum, 'Legal Personhood for Artificial Intelligences' (1992) 70 N.C. L. 1239; Jane C. Ginsburg, 'The Concept of Authorship in Comparative Copyright Law' (2003) 52 DEPAUL 1072.

the article relied on existing legal sources and literature on legal personhood[65] and lifting the corporate veil.[66]

Article II of this research addresses the reidentification of IPR-protected non-personal data and its possible data protection implications. The possibility of using IPR (such as copyright, database protection and trade secrets) and contracts to protect non-personal data has been considered extensively.[67] Hoeren,[68] Zech,[69] and Rees[70] have also made various propositions for the creation of the rights of the data producer, which will amount to the creation of *sui generis* rights for the protection of eligible (non-personal) data elements. The question yet to receive much attention is the status of the right to data protection if previously thought to be non-personal data is processed in a manner that leads to the identification of natural persons. Article II of this research and other parts of this synthesis attempt to resolve this issue by making relevant propositions such as adopting an ad-hoc data producer, reducing reliance on anonymisation as a data deidentification technique, redefinition of personal data, etc.

In considering the data protection issues inherent in using autonomous vehicles in Article III of this research, Hermann et al. considered the history, nature, form and some privacy concerns posed by autonomous vehicles.[71] Maurer et al. also consider various legal and technical aspects of the function of autonomous cars.[72] The EU Handbook on data protection details various principles of data protection law, which apply in different forms to the rights of data subjects whose personal data will be processed by or as a consequence of the use of autonomous vehicles.[73] In highlighting

65 Visa A. J. Kurki, A Theory of Legal Personhood (OUP, 2019); Jessica Berg, 'Of Elephants and Embryos: A Proposed Framework for Legal Personhood' (2007) 59 Hastings L.J. 369; Charlotte O'brien, 'I trade, therefore I am: Legal personhood in the European Union' (2013) 50 Common Market Law Review, (6) 1643–1684.

66 Mayson, French and Ryan, Company law (OUP, 2016) 133-142; Chrispas Nyombi, 'Lifting the Veil of Incorporation Under Common Law and Statute' (2014) International Journal of Law & Management Vol 56 Issue 1, 6-81; Lawrence B. Solum, (n 64) 1239.

67 Taina E. Pihlajarinne and Rosa M. Ballardini, 'Owning Data via Intellectual Property Rights: Reality or Chiemera?' in Rosa Ballardini, Olli Pitkänen and Petri Kuoppamäki (eds), Regulating Industrial Internet through IPR, Data Protection and Competition Law (Kluwer, 2019) 115-133.

68 Thomas Hoeren, 'Big Data and the Ownership in Data: Recent Developments in Europe' (2014) 36 EIPR 751.

69 Herbert Zech, 'A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data' (2016) Journal of Intellectual Property Law & Practice, Vol. 11, 463-464.

70 Christopher Rees, 'Who Owns Our Data?' (SSRN, 2013) <https://ssrn.com/abstract=2310662> accessed 05/08/2020.

71 Andreas Herrmann, Walter Brenner and Rupert Stadler (n 16).

72 Markus Maurer, J. Christian Gerdes, Barbara Lenz and Hermann Winner (eds), Autonomous Driving: Technical, Legal and Social aspects (Springer, 2015).

73 COE/FRA, Handbook on European Data Protection Law, (Publications office of the European Union, 2018). <https://fra.europa.eu/en/publication/2018/handbook-european-data-protection-law-2018-edition> accessed 12/02/2021.

the data protection risks inherent in medical AI systems, Obermeyer's discovery of systemic bias being perpetrated by AI used for medical purposes is also examined in article IV of this research.[74] The key research question in Articles III and IV revolves around how selected AI systems can be deployed in a data protection-compliant manner. This research takes advantage of these existing research publications in addressing various relevant issues (some of which are highlighted above) with the aid of the doctrinal research method. The definition of AI is critical to this research. One of the methods that support the practical identification of AI is the Turing test which lays down the criteria for determining if a computer system is intelligent and is relied upon compared to human intelligence standards.[75] Kaplan's definition of AI and that of Russel and Novig have also been relied upon in the various articles that make up this research. The crux of these definitions is that AI systems process data in an intelligent manner with varying requirements of human intervention.[76]

While there are scholarly works which jointly consider data protection law and IPR, none of those considerations falls within the scope of this research.[77] While there is a vast amount of literature on data reidentification, none of these articles considers reidentification within the context of the reidentification of IPR-protected data.[78] In the conflation of the articles above, the identified vacuum in regulation and scholarly research lies in the data protection risks that might arise in the reidentification of IPR-protected non-personal AI-generated data. This is where the novelty which essentially justifies this research lies. To fill this vacuum, this synthesis proposes the (re)definition of personal data reduce the scope of its current definition, thereby making more data available for use in AI systems. It will also be argued that the current broad nature of the definition of personal data will limit the scope of data available for AI, thereby limiting its development. For this purpose, the works of other scholars who have proposed the redefinition of personal data (based on a different rationale) will also be considered.[79]

---

74 Ziad Obermeyer, Brian Powers, Christine Vogeli and Sendhil Mullainathan, 'Dissecting racial bias in an algorithm used to manage the health of populations' 366 (2019) Science 447-453.

75 Alan Turing (n 7).

76 Kaplan's definition of AI as well as that of Russel and Novig have also been relied upon in the various articles which make up this research. Jerry Kaplan (n 8) 68. Peter Norvig and Stuart Jonathan Russell (n 6) 7.

77 Banterle, Francesco, 'The Interface between Data Protection and IP law: The Case of Trade Secrets and Database Sui Generis Right in Marketing Operations, and the Ownership of Raw Data in Big Data Analysis' in Mor Bakhoum, Beatriz Conde Gallego, Mark-Oliver Mackenrodt, Gintarė Surblytė-Namavičienė (eds), Personal Data in Competition, Consumer Protection and Intellectual Property Law Towards a Holistic Approach? (SSRN, 2016) <https://ssrn.com/abstract=3276710> accessed 24/06/2022. See also Jonathan Zittrain, 'What the Publisher Can Teach the Patient: Intellectual Property and Privacy in an Era of Trusted Privication' (SSRN, 2000). <https://ssrn.com/abstract=214468> accessed 24/06/2022.

78 See Article II of this research.

79 Ibid.

2.1.1 Current Legal Framework

In addition to the above, the EU has adopted various legal instruments as part of its AI regulatory efforts. Though some of these instruments might not particularly apply to data protection and IP law, they remain relevant to some of the AI-related issues addressed in this research. Some of the EU's AI regulatory efforts include the AI High-Level Expert Group (HLEG)'s Guidelines for Trustworthy AI.[80] One notable contribution of the HLEG Guidelines is the seven essential requirements identified as necessary for the legitimate use of AI. These essential requirements are human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental well-being, and accountability.[81] These Guidelines also played a considerable role in the European Parliament's communication on AI.[82] More recently, the EU parliament proposed a Regulation for the governance of AI within the Union.[83] These legal instruments are considered in this research. Besides the AI-specific legal instruments above, the EU parliament also recently published some proposals for (principally, non-personal) data regulations, such as the Data Act and the Data Governance Act, which are subsequently considered in this synthesis.[84]

Other traditional legal instruments also have varying levels of relevance to the issues addressed in this research. To this end, this research considers existing data protection and IPR laws. Some global IPR legislations, such as the Berne Convention,[85] WIPO Copyright Treaty,[86] and the TRIPs Agreement,[87] were considered from the perspective of international copyright treaties. In addition, EU IPR laws such as the Computer Directive,[88] 'Directive on the Information

80 AI HLEG, 'Ethics Guidelines for Trustworthy AI' (EU Commission website, 2019) < https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai > accessed 06/04/2021.

81 Ibid.

82 Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Building Trust in Human Centric Artificial Intelligence (COM(2019)168).

83 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 {SEC(2021) 167final} - {SWD(2021) 84final} - {SWD(2021) 85 final}. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206> accessed 21-07-2021.

84 Footnote 59.

85 Berne Convention for the Protection of Literary and Artistic Works, September 9, 1886, as revised at Stockholm on July 14, 1967, 828 U.N.T.S. 221.

86 WIPO Copyright Treaty, Dec. 20, 1996, S. Treaty Doc. No. 105-17 (1997); 2186 U.N.T.S. 121; 36 I.L.M. 65 (1997).

87 Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994.
Marrakesh Agreement Establishing the World Trade Organisation, Annex 1C, 1869 U.N.T.S. 3; 33 I.L.M. 1197 (1994).

88 Directive 2009/24/EC of the European Parliament and of the council of 23 April 2009 on the legal protection of computer programs (2009) OJ L111.

Society'[89] and the 'Directive on the digital single market' were also considered in addressing the relevant IPR issues in this research.[90] The issues addressed under these global, regional, and national laws include the definition of 'author' under applicable IPR legislation as well as the compliance of AI with principles of data protection.

Relevant judicial decisions have also been analysed in this research. Notable decisions of the Court of Justice of the European Union (CJEU) have been considered where necessary to underline relevant issues in this research. From an IPR perspective, some of such decisions include the notorious decision of the CJEU in the Infopaq case, where the court held that to qualify for copyright protection, a work must be of the author's intellectual creation. [91] Another crucial decision of the CJEU was the Murphy case, where the court held that to meet the originality eligibility of copyright, an author must stamp his personal touch and express his creative abilities in the work.[92] These cases were critical in establishing the court's jurisprudence while making a case for copyright protection for (selected) AI systems. Some of the data protection-related cases which are relevant and cited in this research include the Telesverige case, in which the CJEU held, among other things, that it is unlawful to process personal data when the purpose of processing had not arisen at the point of personal data collection.[93] Another relevant decision is that of the Swiss federal supreme court in the case of EDÖB v Google, where the court held (among other things) that in Google's collection of personal data for the street view, notice ought to be provided to data subjects in both the local and regional media.[94] Both cases were used to espouse the applicability of AI to the principles of data protection law to AI.

89 Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society.

90 Directive (Eu) 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC.

91 In that case, the court held inter alia that to qualify for copyright protection, copyrightable work must be an author's own intellectual creation. Case C-145/10 Eva-Maria Painer v Standard Verlages GmbH et al. [2013] ECLI:EU:C:2013:138, 42.

92 Football Association Premier League Ltd and Others v QC Leisure and Others (C-403/08) and Karen Murphy v Media Protection Services Ltd (C-429/08).

93 Tele2 Sverige AB v Post- och telestyrelsen (C-203/15) EU:C:2016:970; [2017] Q.B. 771, Para 90, 102, 103, 108-110.

94 BGE 138 II 346.

3 SUMMARY OF THE ARTICLES

This section summarises the overall objectives and main contributions of the individual articles within this research.

3.1 Article I

AI-Generated Works and Copyright Law: Towards A Union of Strange Bedfellows - Journal of Intellectual Property Law & Practice*, Volume 16, Issue 2, (2021) 124–135. https://doi.org/10.1093/jiplp/jpaa18*

Abstract

- The creation of copyrightable works was exclusively reserved for humans. However, other non-humans now create copyrightable works.
- Despite this fact, a substantial number of copyright legislations in the European Union (EU) member states do not provide copyright protection for (Artificial Intelligence) AI-generated works.
- Focusing on the authorship eligibility requirement of EU copyright law, this article makes a case for the extension of copyright protection to AI-generated works through the attribution of legal personhood to AI systems that create copyrightable works.
- This article also proposes for, (among other things), the importation of the corporate law principle of 'lifting the veil of incorporation' into the ascription of legal personhood to AI systems.
- The impact of these propositions on AI-generated works and copyright law as well as some probable consequences of not adopting these propositions are also addressed.

3.1.1 Overall Objectives and Main Contributions

This article focuses on the IP protection of AI-generated works. Though it is possible to relate some of the considerations to IP rights generally, the focus of this article is the eligibility of AI-generated works to be protected by copyright law. One of the preliminary findings of this article is that AI-generated works would have been eligible for copyright protection had they been created or authored by natural persons. However, several challenges confront the recognition of AI as an author.

One such challenge is that AI requires varying levels of input from natural persons, which questions the extent to which AI-generated works truly deserve copyright protection.[95] This input requirement implies that AI can be used as a tool controlled by a human person during the creative process.[96] However, this argument does not apply to AI systems that create works independently of human input.[97] Another perspective might be viewing AI-generated works as works of joint authorship between AI and a human author. Nevertheless, this approach still poses some concerns about recognising AI as an author. This article finds a lack of clarity on the definition of an 'author' under international copyright treaties, particularly the Berne convention, as both natural and legal persons can be recognised as authors. This lack of clarity is also reflected in the EU's various (sectoral) copyright legislation, where either a natural and/or legal person might be recognised as an author.[98] However, a cursory look at relevant decisions of the CJEU suggests that only natural persons can be recognised as authors.[99] Amongst EU member states, there is evidence that both natural and legal persons might be recognised as authors. For instance, this is the case in Ireland, where section 2 Copyright and Related Rights Act 2000 defines computer-generated works as works 'generated by computers in circumstances where the author of the work is not an individual'.

The challenges identified above are active threats to the development of AI and AI-generated works and ought to be resolved if AI and its works are to gain global traction. To resolve these challenges, the authorship eligibility requirements must be revisited under the jurisprudence of EU law. One of the primary propositions made by this article is the call for the ascription of legal personhood to AI systems to enable the recognition of AI systems as authors of their copyrightable works. However, this proposition faces many roadblocks, paramount among which is assigning moral rights to AI-generated works. This challenge is more apparent in civil law jurisdictions where moral rights seek to protect the author's personality and 'subjective feelings' even beyond economic rights.[100] Some features of the civil law approach to moral rights include, among other things, the author's right to the integrity of the work, the right to prevent an unlicensed modification of the work,

95 Pamela Samuelson, (n 63) 22.

96 Jane C. Ginsburg, (n 64) 1072.

97 Ahmed Elgammal, (n 28).

98 Art 4(1) Directive 96/9/EC of the European Parliament and of the council of 11 March 1996 on the legal protection of databases [1996] OJ L 77/20.

99 Case C-145/10 Eva-Maria Painer v Standard Verlages GmbH et al. [2013] ECLI:EU:C:2013:138, 42; Case C-145/10 Eva-Maria Painer v. Standard Verlages GmbH etal. [2013]ECLI:EU:C:2013:138. 42, Ground 2 of the decision; Case C-604/10 Football Dataco Ltd etal., v Yahoo! Etal [2012] ECLI:EU:C:2012:115.

100 Adam D. Moore and Kenneth Einar Himma, 'Intellectual Property' in Edward N. Zalta (ed), (Stanford Encyclopedia Of Philosophy, 2011) paragraph 2.7 <https://ssrn.com/abstract=1980917> accessed 01/05/2021.

the right to be recognised as the author of the work, etc.[101] This human-centric moral rights approach poses a challenge to AI systems because AI lacks the human nature to have moral rights vested. In this article, a case is made for the proposed ascription of legal personhood to AI systems. The fact that AI is intended to futuristically simulate human intelligence[102] and the outright scrapping of moral rights are some of the arguments made to justify the ascription of legal personhood to AI.[103] One of the significant constraints to ascribing legal personhood to AI is the EU's technology neutrality approach to regulating IPR. The technology neutrality approach requires that IPR legislation be drafted to ensure it focuses on the nature and use of the work rather than the medium for its creation.[104] The rationale behind this regulatory approach includes securing the validity and longevity of legislation irrespective of technological changes and its online and offline enforceability.[105] However, this article downplays the importance of this approach for reasons which include derogating some existing laws from it.[106] Furthermore, the article argues that the current practice of negotiating appropriate borders for technology neutrality in legislation rather than adequately regulating them might be inappropriate owing to the undeniable global impact of technology.

A general proposition that has been made in the regulation of AI-generated works is that such works should lie in the public domain. Article I considers the pros and cons of this proposition because an article of this nature will be incomplete if such relevant divergent views are not addressed. One of the arguments in support of the 'public domain' proposition is that creatives are not incentivised by financial gains but by fame, passion, and recognition.[107] The article counters this position by suggesting that fame, passion, and recognition might not be sufficient motivating factors, especially when developers have built AI systems for 'creating artistic works in certain specific ways that, for instance, cannot be done by human beings themselves' or in cases where the AI system itself would not be eligible for IP protection but the work created by the AI system would.[108] The article further argues that leaving AI-generated works in the public domain might devalue human-authored works because of AI's ability to create large volumes of work within a limited time. This

101 Cyrill P. Rigamonti, 'Deconstructing Moral Rights' 47(2) (Harvard International Law Journal, 2006) 362-367.

102 Stuart J. Russell and Peter Norvig (n 6) 7.

103 An example of this is visible in section 178 UK Copyright, Designs and Patents Act 1988 (UK CPDA) provides inter alia that the right to be identified as the author or director of a work shall not apply to computer-generated works.

104 Article 6 WIPO Copyright Treaty.

105 Chris Reed, 'Taking Sides on Technology Neutrality' (Scripted, 2007) Volume 4, Issue 3 <https://script-ed.org/wp-content/uploads/2016/07/4-3-Reed.pdf> accessed 11/07/2021.

106 See section 178 UK CPDA

107 Linda J. Lacey, Of Bread and Roses and Copyrights (1989) Duke Law Journal 1532-1596.

108 Rosa Ballardini, Kan He and Teemu Roos (n 40).

might result in a situation where there will be a natural inclination towards AI-generated works which are freely available in the public domain, thereby leaving the works created by human authors with lesser royalties and patronage. This possible disincentivisation of human authors is a societal risk due to the lack of consciousness in AI systems. This lack of consciousness means that (at the time of writing this synthesis) AI systems are incapable of investigating the causal effects of problems and can only compile or analyse problems.[109] Having taken the competing positions into perspective, the article concludes by calling for an amendment of relevant laws relating to the definition of authorship and the ascription of AI systems with legal personhood.

3.2 Article II

Balancing Competing Interests in the Reidentification of AI-Generated Data - *European Data Protection Law Review Vol 8 Issue 3 (2022) 362 - 376*

Abstract
AI systems generate valuable analytical information from (apparently) non-personal data with vast economic consequences. This information generated from non-personal data provides a competitive edge which serves as a critical rationale for its appropriation to the exclusion of others. The proliferation of AI has made it possible for non-personal data (including anonymised data) to result in the reidentification of natural persons. There have been propositions from various quarters for protecting non-personal data through Intellectual Property Rights (IPR). Should AI-generated data be protected by IPR, this can potentially result in data protection concerns in the event of data reidentification that singles out natural persons. This might particularly occur where reidentified data leads to the identification of natural persons in circumstances where the applicability of data protection law had neither been contemplated nor anticipated. This article highlights the concerns that might arise in the event of data reidentification and how this might raise exciting challenges for data protection compliance.

3.2.1 Overall Objectives and Main Contributions

This article is very central to the theme of the intersection between data protection and IP law. In the operation of AI systems, large volumes of big data (particularly non-personal data) must be processed if such systems are to perform their tasks.

109 Jonathan Zittrain, 'The Hidden Costs of Automated Thinking' (The New Yorker, 23 July 2019). <https://www.newyorker.com/tech/annals-of-technology/the-hidden-costs-of-automated-thinking> accessed 09/05/2021.

Therefore, AI systems generate data which provides analytical information that is an excellent source of competitive advantage for (commercial) organisations and entities. The critical nature of data to the operation of AI systems and its competitive advantage has elevated its status as an asset that AI manufacturers are eager to protect. As a result, AI-generated data are largely protected by contracts, with a strong case being made for its IPR protection.[110] However, any attempt to protect AI-generated data through IPR might pose some data protection concerns. This is because of the enormous possibilities posed by data reidentification which blurs the line between personal and non-personal data. Scholars have undertaken extensive research to prove that non-personal data can be combined with other data sets to identify natural persons.[111] AI also plays a role in the reidentification process.[112] The implication is that should AI-generated data be protected by IPR, data reidentification might result in easily converting such non-personal data to personal data, thereby making data protection applicable therein. This article aims to highlight this risk and propose solutions for its mitigation.

The major contribution of this article lies in its highlighting of potential data protection and IP concerns as well as its proposition of possible solutions. Some of the data protection concerns which might arise from the ascription of IPR to AI-generated data include data reidentification by third-party licensees, the resolution of data subject requests about reidentified data, and the (in)effectiveness of data anonymisation as a tool for excluding the applicability of data protection law. In the case of data reidentification by third-party licensees or any other party having (un) lawful access to the data, there is a huge possibility that these parties might combine the data with other data sets obtained from other sources with the effect of identifying data subjects. Therefore, data reidentification can result in data subject requests for information or other GDPR rights, creating data protection concerns because of the lack of data protection considerations in initiating such cases. These possibilities that abound with data reidentification highlight the limitations of data anonymisation and questions its effectiveness as a tool for excluding the applicability of data protection law. To mitigate the identified concerns in this article, it is proposed to revisit the status of data anonymisation under data protection law. Furthermore, it is suggested in this article that rather than focusing on data anonymisation as a tool for excluding the applicability of data protection law, legal instruments should focus on the prevention of data reidentification. It is further proposed that the possibilities of

110 Péter Mezei, 'From Leonardo to the Next Rembrandt – The Need for AI-Pessimism in the Age of Algorithms' (SSRN, 4 May 2020). < https://ssrn.com/abstract=3592187> accessed 29/09/2021.

111 Latanya Sweeney, 'Matching Known Patients to Health Records in Washington State Data' (SSRN, 5 June 2013) <https://ssrn.com/abstract=2289850> accessed 06/04/2021.

112 Karl Bode, 'Harvard Students Again Show Anonymised' Data Isn't Really Anonymous' (Techdirt, 20 February 2020) <https://www.techdirt.com/articles/20200203/07405543847/harvard-students-again-show-anonymised-data-isnt-really-anonymous.shtml> accessed 06/04/2021.

data reidentification and the effectiveness of the measures taken to prevent it should be communicated to data subjects as part of the requirements of the transparency principle.

The article explicitly calls for balancing competing interests between the rights holder and the data subject regarding data subject rights. This is achievable by ensuring that, where possible, reidentified personal data is distinguishable from the creative elements of data. Privacy by design principles which will make reidentified data distillable from any protected creative elements of data, ought to be introduced to prevent the publication of IP rights in the preceding scenario. The appointment of an ad-hoc data controller responsible for acting as a controller in the event of data reidentification ought also to be considered to protect any arising data protection rights.

3.3 Article III

Autonomous Transport Vehicles vs The Principles of Data Protection Law: Is Compatibility Really an Impossibility? *International Data Privacy Law*, Volume 10, Issue 4, (2020) 330–345, https://doi.org/10.1093/idpl/ipaa017

Abstract

- Autonomous (transport) vehicles have evolved from science fiction into a feature of reality (in) which we now live.
- From a data protection standpoint, one of the challenges confronting the integration of AV into the society is the question of whether or not this disruptive technology is capable of being compliant with the principles of data protection law.
- The importance of focusing on the principles of data protection law lies in the fact that these principles encompass the entire body of data protection law. Failure to achieve compliance with said principles automatically amounts to a failure to comply with data protection law.
- With a focus on the European Union and the European Economic Area, this article seeks to identify the extent to which the extant data protection laws are capable of protecting the right to data protection of data subjects in the use of autonomous vehicles while also making recommendations on how compliance can be best achieved.

3.3.1 Overall Objectives and Main Contributions

This article focuses on the data protection implications that flow from using autonomous vehicles in the European Union/European Economic Area (EU/EEA). Autonomous cars and ships are the main types of autonomous vehicles

that form this article's focus.[113] To appreciate its data protection implications, this article provides some background into the nature and methods of data collection in autonomous vehicles. One clear fact about AI is its heavy reliance on data, and autonomous vehicles are no exception. Autonomous cars collect various categories of (personal) data, including travel habits, destinations, stops, routes, body size, number of passengers, musical taste, environmental data, etc.[114] Autonomous vehicles also possess 'home and emergency buttons', which can either cause passengers to be driven home or connected to ambulance services. Data recorders which can either collect autonomous vehicle crash-related data or data related to the entire journey, also collect data in autonomous vehicles.[115] These (personal) data collected by autonomous vehicles are collected using sensors.[116] Autonomous ships also collect large volumes of (personal) data like autonomous cars, such as images of persons[117] and cargo ID numbers, which might lead to the identification of natural persons, etc. The definition of autonomy is a crucial distinguishing factor between autonomous cars and ships. As anticipated in this article, autonomous cars are entirely autonomous, but the same cannot be said of autonomous ships.[118] Unlike autonomous cars, the classification of a ship as autonomous does not mean there won't be crew members on the ship able to take over the control of the ship in emergencies.[119] Also, autonomous ships will have a Shore Control Centre (SCC) where the traditional functions of the crew will be performed. The SCC conducts

113 Other types of autonomous vehicles which are autonomous trains and autonomous ships are yet to see as much technological development as AC and AS.

114 Kai Rannenberg, 'Opportunities and Risks Associated With Collecting and Making Usable Additional Data' in Markus Maurer, J. Christian Gerdes, Barbara Lenz and Hermann Winner (eds), Autonomous Driving: Technical, Legal and Social Aspects (Springer, 2015) 499-500 <https://link.springer.com/content/pdf/10.1007%2F978-3-662-48847-8.pdf> accessed on 19/06/2021.

115 Andreas Herrmann, Walter Brenner and Rupert Stadler (n 72) 94.

116 Such sensors include radar, lidar, ultrasound, and cameras.

117 This could include the images of persons captured from the surrounding of the ship, or captured in man overboard situations, etc.

118 Five levels of automation have been identified for AC. Level 0 involves no automation at all; in Level 1 and 2, the system takes over some of the driving tasks but the driver is required to continually monitor the system and must be able to take over the driving as soon as it becomes necessary. Level 3 requires less monitoring of the system by the driver while in level 4, the system is able to drive the car in normal operation and in defined surroundings with the driver being able to intervene at will. Level 5 is the final stage, which is the fully automated and autonomous driving stage and is the focus of this article. Please see: Andreas Herrmann, Walter Brenner and Rupert Stadler, (n 16) 8-9, 47-51.

119 Six autonomy levels (AL) have been adopted for AS. These are AL 0 where there is no autonomy at all. AL 1-5 involves different levels of autonomy with varying levels of human monitoring and intervention, while AL 6 involves full autonomy with no human intervention. Mogens Blanke, Michael Henriques and Jakob Bang (n 18) 1, 3-6.

a lot of data processing[120] and the lookout function, which is now performed by sensors.[121]

The crux of this article is the identification of potential data protection concerns which arise from the use of autonomous vehicles. The principles of data protection law are used as a benchmark for identifying these potential concerns. One such potential concern pertains to the lawfulness and transparency principle.[122] One of this article's findings is that using a justifiable legal basis and transparent data processing may not be readily achievable because of the large volumes of data being processed and the unawareness on the part of data subjects (in some cases) about the data collection.[123] As a remedy, enacting specific legislations/amendments of existing legislations (e.g. traffic laws) is a possible solution to this legal basis dilemma. The use of radio jingles, community town hall meetings as well as a diagram of an information notice to be placed on highways are suggested in the article as a solution to the transparency concern.[124] The diagram proposed by this article for the transparent communication of the data processing relevant information is one of the main contributions of this article to the body of knowledge on this topic.[125] Another concern addressed in this article is the data minimisation principle which arises because of the possibility of autonomous vehicles (and AI generally) collecting more personal data than is necessary for the processing activity.[126] Another suggestion of this article is the use of adequate technical and organisational measures such as privacy by design and default,[127] anonymisation[128] and/or pseudonymisation[129] as helpful tools that could help in the minimisation of data collected by the Controller. Also relevant are the confidentiality and integrity principles, which promote the security of personal data,

120 Rolls-Royce has released some videos showing how a SCC is operated and used in controlling autonomous ships. These videos reveal that such SCC will have access to and control all data pertaining to the ship. See: Rolls-Royce, 'Ship Intelligence for Cargo Vessels' (Youtube, December 2014). <https:andandwww.youtube.comandwatch?v=_nApv-C7qSg&list=PLk-17K0buHIvy68TGjnSUppTq-Gi91lT-> aaccessed 27/05/2021.

121 The Norwegian Forum for Autonomous Ships (NFAS), 'Definitions for Autonomous Merchant Ships' in Ornulf Jan Rodseth and Hävard Nordahl (eds), (NFAS website, 2017) 7, 16-18. <https://nfas.autonomous-ship.org/wp-content/uploads/2020/09/autonom-defs.pdf> accessed 19/06/2021.

122 Recital 40 GDPR, Art 6 (1) (a)-(f) GDPR.

123 The Information Commissioner's Office, (n 5) 12.

124 The use of public and private media for meeting transparency requirements has found support in EDÖB v Google, BGE 138 II 346.

125 Emmanuel Salami, 'Autonomous transport vehicles versus the principles of data protection law: is compatibility really an impossibility?' (November 2020) International Data Privacy Law, Vol 10, Issue 4, 342.

126 Nikolaus Forgó, Stefanie Hänold, and Benjamin Schütze, 'The principle of big purpose limitation and big data' in Marcelo Corrales et al. (eds.), New Technology, big data and the law (Springer, 2017) 20.

127 Privacy by design requires the incorporation of data protection principles into data processing activities at the design phase. Art 25 GDPR. Recital 78 of the GDPR.

128 Recital 26 GDPR.

129 Art 4(5) GDPR.

including protection against unauthorised or unlawful processing and accidental loss, destruction, or damage.[130] As a result of the large volumes of data processed by multiple processors in autonomous vehicles, these principles are geared towards ensuring the security of data against the activities of unauthorised persons who may unlawfully access the data. Other data protection principles, such as the fairness principle, purpose limitation principle, and storage limitation, are also considered. In conclusion, this article finds that autonomous vehicles can comply with data protection law if relevant principles are considered at the design and deployment stages.

3.4 Article IV

AI, Big Data and The Protection of Personal Data in Medical Practice. E*uropean Pharmaceutical Law Review, Vol 3 Issue 4 (2019) 165 - 175*

Abstract

The introduction and use of AI and big data is fast becoming a norm across various sectors of the global economy and the health sector is no exception. As of today, AI is being deployed for disease diagnosis, clinical trial research, drug discovery, medical consultations to mention just a few. The implication of the usage of AI in this manner is that large amounts of personal data will be processed by AI systems designed for use in the medical space. This article seeks to examine the various uses of AI in medical practice within the European Union/European Economic Area with the objective being the analysis of potential privacy challenges that may arise therefrom.

3.4.1 Overall Objectives and Main Contributions

AI systems are now critical participants in the medical sector, rendering various services fundamental to maintaining life and preserving well-being.[131] However, this has not been without data protection implications. This article examines the effects of these implications within the medical space while making appropriate recommendations. One key factor necessitating the consideration of this topic is the sensitive nature of the data being processed, as well as the involvement of vulnerable and sick people as data subjects.[132]

130 Art 5(1) (f) GDPR, Article 32 GDPR.

131 Some of the medical services provided by AI include disease diagnosis, drug discovery and repurposing, telerobotic surgeries and robot-assisted surgeries, genomics, etc.

132 Art 9 GDPR.

One of the data protection concerns identified in this article relates to decision-making algorithms that might affect the rights and freedoms of data subjects.[133] One of the ways through which AI systems could make decisions that affect individuals' rights and liberties is when AI systems diagnose patients with ailments. In applying diagnoses, AI systems could play a role in determining a person's employability.[134] Therefore, data subjects ought to be provided with adequate information about the processing activity and the considerations used by AI in making its decisions for AI systems to comply with data protection law.[135] There is also the need to have such a decision reviewed by a natural person to determine its correctness and potentially rectify any errors therein.[136] However, the major problem remains the opaque nature of today's AI algorithms which makes an explanation of the rationale behind its decision-making almost impossible.[137]

Furthermore, data security concerns in using AI systems for medical purposes are heightened because of the large volumes of sensitive personal data processed therein. These large volumes of sensitive personal data would potentially pass through a large group of (non-medical) personnel that would ordinarily have had no access to it in traditional medical data processing activities. This could occur through sharing such data across multiple jurisdictions, uploading them on various systems, the performance of system updates and related services by engineers and other security experts, etc. Data security poses a potential risk in using AI in medicine if not adequately addressed. The data security concern also includes the risk of possibly transferring personal data outside the EU/EEA. This is particularly true with telerobotic surgeries, typically carried out across multiple jurisdictions. In such cases, EU rules for data transfers out of the union ought to be complied with.[138]

Another data protection concern identified in this article is the possibility of AI meting out discrimination on data subjects in the provision of medical services. This could be the direct result of using biased data in AI systems resulting in subjective outcomes which deny people medical rights. An example of this occurred in the USA, where people of colour were unfairly denied access to healthcare when compared to their Caucasian counterparts as a result of discriminatory bias.[139] Though the article acknowledges that almost every decision-making process involves discrimination in some form, what the law opposes are unlawful grounds for discrimination. This

133 Art 22 (1) GDPR.

134 AI systems are being used to diagnose heart conditions. See Michael Quartermain, 'Echocardiogram' in E. Alboliras, Z. M. Hijazi, L. Lopez and D. J. Hagler (eds), Visual Guide to Neonatal Cardiology (Wiley Online Library, 2018) <doi:10.1002/9781118635520.ch13> accessed 01/08/2021.

135 Art 13 GDPR.

136 Recital 71 and Art 22 GDPR.

137 Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan, Cass Sunstein, 'Discrimination in the Age of Algorithms' (SSRN, February 5, 2019) <https://ssrn.com/abstract=3329669> accessed at 10/12/2021.

138 chapter 5 GDPR.

139 Ziad Obermeyer, Brian Powers, Christine Vogeli and Sendhil Mullainathan (n 74) 447-453.

means that any decision-making entity (in this case, AI) necessarily discriminates as a matter of course in the sense of making distinctions between people based on specific (lawful) features. However, it is the act of making such distinctions between people based on unlawful characteristics such as sex, race, colour, language, religion, political opinion, etc., that the law frowns upon.[140]

In resolving these concerns, some solutions specific to the use of AI in medicine are proposed. One measure which can go a long way in mitigating some of the risks identified in this article is the privacy by design principle, which requires incorporating the privacy principles at the technology development stage. The article also suggests standardising the security measures used in medical AI devices. This can be achieved through a data protection certification body which will design the necessary safeguards subject to the relevant provisions of the GDPR.[141] Furthermore, the article suggests using codes of conduct to standardise anti-discrimination practices in the development of AI. Such codes of conduct would ensure that data entered into AI systems for processing is diverse and that AI developers are representative of diverse ethnic, cultural, and economic backgrounds. AI systems would also benefit from audits from researchers and other industry stakeholders who might be able to identify possible violations of privacy principles.[142] Finally, the article also advocates for better standardisation of the practice of AI developers in medicine to ensure better accountability. This is because, ordinarily, medical doctors who treat patients are well-trained. Therefore, there is no reason why AI medical developers should also not be made to undergo standardised training. Such standardisation will also foster the ethical standards between doctors and patients in the case of medical AI developers and relevant data subjects.[143]

140 Art 14(1) European Convention on Human Rights.

141 Art 42 and 43 GDPR

142 For instance, this is how the violations in the USA were identified.

143 Brent Mittelstadt, 'Principles Alone Cannot Guarantee Ethical AI' (Nature Machine Intelligence, November 2019). <https://ssrn.com/abstract=3391293> accessed 22/04/2021

4 RISKS AND (THE IMPLICATIONS OF) REGULATORY INTERVENTION

The totality of this research (i.e. the four articles and this synthesis) identifies potential risks within the fields of data protection and IP law which can (concurrently) arise in the use and adoption of AI. As specified in earlier parts of this research, some of these risks can occur in the (non-) ascription of IPR to AI systems, the reidentification of AI-generated non-personal data, data protection risks in the use of AI systems, etc. This section uses some of the data protection risks that might arise during data reidentification to emphasise the criticality of the risks identified in this research. More specifically, the possibility of data subjects exercising their access rights in the event of the reidentification of non-personal data is used to embody the severity of the risks identified in this research.

Furthermore, the effect of regulatory intervention is considered through selected recent legislative proposals of the EU. Some of these legislative proposals were yet to be issued when this research article was published. These legislative proposals might have two possible consequences - they might either play a role in resolving the risks identified above or highlight these risks.

4.1 Data Subject Rights: A Brief Exposé on the Risks Justifying this Research

This research contends that data protection rights might be infringed upon if IPR-protected non-personal data is reidentified in a manner that singles out natural persons. This section highlights the risk identified in this research, particularly regarding data subject rights. Data subject rights are requests from data subjects directed at a data controller in the exercise of their rights. Such rights include the right to data access, rectification, deletion, portability, etc.[144] One of the key contentions of this research is that any ascription of IPR to AI-generated data can potentially affect data subjects, particularly in exercising their access rights. In such an event, non-personal data protected by IPR might be reidentified, thereby making it personal data and subject to data protection law. The right to data portability is one of the access rights which entitles data subjects to receive personal data concerning them, which they have provided to a data controller in a structured, commonly used,

144 See Chapter 3 (Art 12-23) GDPR.

and machine-readable format and have the right to transmit said data to another controller without hindrance.[145]

Should AI-generated non-personal data be protected by IPR, technology and other reidentification techniques can create the possibility of data reidentification, thereby converting previously (believed to be) non-personal into personal data.[146] In the event of such reidentification, data subjects will be entitled to exercise their access rights in a situation where there is neither a justifiable legal basis preceding the commencement of the processing activity nor a designated data Controller responsible for exercising the roles attributed with that designation under the GDPR.[147] This is further aggravated by the fact that data subjects could make data requests from data controllers even in situations where the relevant data forms a part of protected works. In such a scenario, a data subject exercising their access rights might request the deletion of their data, which has already formed part of the IPR of a rightsholder. This will therefore require some form of balancing the competing interests arising in data protection and IP law. It appears that the CJEU favours the balancing of competing interests between the relevant fields of law.[148] However, this research strives to anticipate and prevent the occurrence of this possibility, however futuristic and hypothetical it may seem. Further analysis of the consequences flowing from the above scenarios will also be considered in the parts of this research where legal empiricism is used. This research will subsequently propose the redefinition of the scope of personal data as a possible solution for resolving the issues identified above.

4.2 Effects of Recent Regulatory Efforts

Following its acknowledgement of the critical nature of AI,[149] the EU, through its legislative apparatus, has made some regulatory efforts towards the governance of AI. Expectedly, these regulatory efforts have been received differently, with some

145 Art 20 GDPR. To be eligible for this right, 'personal data' ought to be processed by automated means, or pursuant to consent or contract and be 'provided by' the data subject. The rationale behind this right has been traced to the promotion of competition by potentially lowering barriers to digital markets entry in the EU. Orla Lynskey, 'Article 20. Right to data portability' in Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), The EU General Data Protection Regulation (GDPR): A Commentary (2020) 497-507.

146 Karl Bode, (n 112).

147 Art 24 GDPR.

148 C-461/10, Bonnier Audio AB, Earbooks AB, Norstedts Förlagsgrupp AB, Piratförlaget AB, Storyside AB v. Perfect Communication Sweden AB, 19 April 2012.

149 White Paper On Artificial Intelligence—A European approach to excellence and trust COM(2020) 65 final <https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf> Accessed 09/06/2021.

stakeholders either approving or disapproving of them.[150] Some of the regulatory efforts of the EU include the Proposal for the Data Act, the Proposal for the AI Act, and the Proposal for the Data Governance Act.[151] For completeness, the impact of the Proposal for the Data Act and the Proposal for the AI Act are mainly considered in relation to this research. However, these two proposals have been selected from several relevant proposed legal instruments because they sufficiently illustrate the interrelationship between data protection and IP law, which this research seeks to highlight.[152]

4.2.1 Proposal for the Data Act

The proposal for the Data Act was adopted on February 23 2022, to harmonise the EU rules on fair access to and use of data. The Proposal is an offshoot of the EU Parliament's request of the EU Commission to propose a Regulation for non-personal data governance in the EU. This Proposal is particularly relevant to this research because of its adoption to harmonise the framework specifying who, other than the manufacturer or data holder, is entitled to access data generated by products or related services as well as the conditions and basis for doing same.[153] It's essential to consider its impact since the proposal is relatively recent and was adopted mainly after the publication of the articles or their acceptance for publication. The Proposal is consistent with the existing laws for protecting personal data (i.e. the GDPR), the rules for preserving the right to privacy, and any private and non-personal data stored in and accessed from terminal equipment (the ePrivacy Directive or any laws that might repeal it).[154] The Proposal applies to products such as vehicles, medical and health devices, etc. While the data generated by these products fall within the scope of the Proposal, the information derived or inferred from the data does not.[155]

---

150 Luciano Floridi, 'The European Legislation on AI: a Brief Analysis of its Philosophical Approach' (2021) 34 Philos. Technol. 215–222 (2021) <https://doi.org/10.1007/s13347-021-00460-9> accessed 09/12/2021. See also Human Rights Watch, 'How the EU's Flawed Artificial Intelligence Regulation Endangers the Social Safety Net: Questions and Answers' (Human Rights Watch, November 2021). <https://www.hrw.org/news/2021/11/10/how-eus-flawed-artificial-intelligence-regulation-endangers-social-safety-net> accessed 09/09/2021.

151 Footnote 58. Some other regulatory efforts of the EU that are relevant to this research are itemised in: Deloitte, 'Artificial Intelligence Act, May 2021 Risk Advisory' 3. <https://www2.deloitte.com/content/dam/Deloitte/de/Documents/Innovation/Deloitte-TAI-DE-Artificial-Intelligence-Act.pdf> accessed 09/09/2021.

152 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 {SEC(2021) 167 Final} - {SWD(2021) 84 Final} - {SWD(2021) 85 Final}2021/0106 (COD).

153 Recital 4 Proposal for the Data Act.

154 Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) OJ L 201, 31.7.2002. See also paragraph 1 of the explanatory memorandum to the Proposal.

155 Recital 14 Proposal for the Data Act.

Before concluding a contract for the purchase, rent, or lease of a product or the provision of a related service, the user of such a product should be provided with clear and sufficient information about how data generated in such a product can be accessed.[156] This disclosure obligation seeks to ensure that users of such products, irrespective of whether they are natural or legal persons, can access the data generated from their use of the product. In cases where the user cannot directly access data from the product, the data holder shall make such data available to the user.[157] In making these disclosures, existing IP rights must be protected.[158]

The Proposal introduces rules for regulating the transfer of non-personal data in the EU. It states that providers of data processing services shall take all reasonable technical, legal, and organisational measures, including contractual arrangements, to prevent the international transfer or governmental access to non-personal data held in the Union where such transfer or access would create a conflict with Union law or the national law of the relevant Member State.[159] Furthermore, the decisions of a court, tribunal or administrative authority of a third country requiring a provider of data processing services to transfer from or give access to non-personal data shall only be recognised if subject to an internal agreement between that third country and the EU or a member state.[160]

4.2.1.1 What's the impact of the Proposal for the Data Act on this Research?

The proposal impacts this research in several ways. First, even though the objective of the regulation is to ensure the free flow of non-personal data, it (at least) incidentally acknowledges the blurred lines between both personal and non-personal data. This can be observed from the proposal's requirement that the principles of data minimisation and data protection by design and by default ought to be adopted when there might be significant risks to the fundamental human rights of individuals. This importation of data protection law principles into the processing of non-personal data can be inferred as an acknowledgement by the European Commission (EC) of the blurred lines between personal and non-personal data. As explicitly stated in the Proposal, the principle of privacy by design and default is relevant when the (non-personal) data processing activity can pose significant risks to the rights of individuals.[161]

Further evidence of the Proposal's acknowledgement of the relationship between personal and non-personal data, particularly regarding the reidentification of personal data, can be gleaned from the provisions regulating the transfer of non-personal

156 Recital 23, Article 3 Proposal for the Data Act.
157 Art 4 Proposal for the Data Act.
158 Art 4(3) Proposal for the Data Act.
159 Art 27 Proposal for the Data Act.
160 Art 27(2) Proposal for the Data Act.
161 Recital 8 Proposal for the Data Act.

data. The Proposal introduces new rules by which providers of processing services shall take all reasonable technical, legal, and organisational measures, including contractual arrangements to prevent the international transfer or governmental access to non-personal data held in the EU where such transfer or access would create a conflict with EU or member state law. By this provision, the proposal imports the data security principle of data protection law into the processing of personal data.[162] This approach will be perfect for protecting non-personal data, mainly in the event of data reidentification.

The Proposal is also relevant to the IPR component of this research. One relevant evidence of this is the requirement that adequate measures for protecting trade secrets should be considered before transferring any relevant data.[163] Therefore, it can be anticipated that when the trade secret of a natural/legal person can be deduced from data transfer, modalities for protecting the trade secret must be entered into between the parties even while also ensuring the free flow of such data. Another interesting consideration is that while data generated from any relevant products (such as vehicles, medical and health devices, etc.) fall within the scope of this research, the information generated from such products does not. This means that the Proposal does not seek to compel the free transfer of information generated from data by the efforts and investment of another person or entity.[164]

4.2.2 Proposal for the AI Act

On April 21, 2021, the EC published a proposal for the AI Act.[165] Even though the Act is only a proposal at the time of writing, research of this nature which is focused on the use of AI within the EU, will be incomplete without its consideration. The AI Act adopts a risk-based approach that regulates AI based on its potential risk(s) not to stifle innovation.[166] The Proposal for the AI Act establishes a relationship with the IP law component of this research by expressly mandating compliance with relevant IPR, particularly the protection of trade secrets, in the use of AI systems. For instance, transparency obligations under the Proposal ought not disproportionately to affect the right to the protection of IPR guaranteed under EU law.[167] One reasonable inference that can be drawn from the Proposal is its acknowledgement of the relationship between IPR and data protection law, thereby

162 Art 32 GDPR.

163 Art 4(3) and 5(8) Proposal for the Data Act.

164 Recital 14 Proposal for the Data Act.

165 For further reading and analysis of the proposal, see Michael Veale, Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act — Analysing the good, the bad, and the unclear elements of the proposed approach'(2021) Computer Law Review International vol. 22, no. 4, 97-112. <https://doi.org/10.9785/cri-2021-220402>

166 Paragraph 2.3 of the Proposal for the AI Act.

167 See paragraph 3.5, (explanatory memorandum), Proposal for the AI Act. See also Article 70 of the Proposal for the AI Act.

aligning with the balancing of competing interests between both fields of law as proposed in Article II of this research.[168] The proposal for the AI Act uses the risk-based approach by outlining prohibited risks and risks that require stringent and less stringent compliance measures depending on their severity. Therefore, the extent of compliance required is determined by the potential severity of the risk posed by the AI. The Proposal also advances a human-centric approach to AI regulation in the EU so that people can be assured that AI is used in compliance with their rights, an approach similar to the objectives of this research.

The proposal for the AI Act adopts an expansive definition of AI systems by defining it as software developed with systems such as machine learning, expert and logic systems, and Bayesian or statistical approaches.[169] As defined under the Act, such AI systems can, subject to human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments with which they interact.[170] The AI Act also categorises AI into three groups the prohibited group,[171] the high-risk group,[172] and a third group classified through specified use cases. This synthesis will refer to this third group as the AI systems with limited risk.[173] The prohibited group contains those use cases of AI that are expressly banned under the Proposal.

Article 56 of the AI Act establishes the European Artificial Intelligence Board (the Board). The Board is charged with the responsibility of contributing to the effective cooperation of the national supervisory authorities and the EC concerning the provisions of the AI Act; coordinating and contributing to guidance and analysis by the EC and the national supervisory authorities and other competent authorities on emerging issues across the internal market concerning the AI Act; and assisting the national supervisory authorities and the EC in ensuring the consistent application of the AI Act.[174] The Board shall be composed of representatives of EU member states and the EC. Member states are also required to set up national competent authorities for enforcing the provisions of the AI Act.[175]

4.2.2.1 What's the impact of the Proposal for the AI Act on this Research?

The Proposal is undoubtedly relevant to this research's data protection and IP law issues. The definition of AI systems adopted under the Act[176] correlates with

168 See Paragraph 3.5 (explanatory memorandum), Proposal for the AI Act.
169 Annex I AI Act.
170 Article 3 (1) AI Act.
171 Art 5 AI Act.
172 Some topics (such as aviation, cars, medical devices.) pertaining to the use of AI are harmonised under this provision. See Annex II of the AI Act.
173 Title IV, Article 52 Proposal of the AI Act.
174 Art 56 (2) (a)-(c) AI Act.
175 Art 59 AI Act.
176 Article 3 (1) AI Act.

the definition adopted in the various articles making up this research, particularly regarding AI interacting with the environment and the requirement of human input.[177] One point of divergence in the definition of AI between the proposal and this research is that while the proposal takes a narrower approach by referring to AI as software, this research takes a broader approach by referring to AI as a process and, in some cases, as a system. The GDPR is expected to apply to AI systems and complement the AI Act as anticipated in this research.[178]

Another relevant impact of the Proposal is the proposed establishment of the European AI Board, which will be set up to operate similarly to the European Data Protection Board (EDPB).[179] The Board will be to AI regulation, what the EDPB is to data protection regulation. In other words, the AI Board will serve as a watchdog responsible for enforcing compliance with the Proposal. Coincidentally, Article IV of this research proposed that establishing a regulatory body was necessary for monitoring, mitigating, and preventing some of the risks identified therein.[180] One fundamental role that the Board will play relates to its power to issue guidance documents on contemporary and emerging AI (related) topics. This benefits the AI Board with ongoing regulatory authority to monitor AI-related developments and issue guidance.

It can be inferred that the Proposal anticipates some level of cooperation between the AI Board and the respective data protection supervisory authorities and/or the EDPB created under the GDPR.[181] This is because the Proposal anticipates some form of consistency between it and the GDPR, particularly as it pertains to high-risk AI systems.[182] To achieve this consistency, one can expect some form of cooperation between the AI Board and national data protection supervisory authorities/the EDPB. This will then result in harmonisations in applying EU data protection laws and the Proposal to AI, limiting the possible risk(s) identified in this research.[183] Furthermore, the Proposal furthers the regulation of AI and reduces the role of ethics therein. Before the introduction of the Proposal, it was common for scholarly works to propose ethical considerations as a possible solution for respecting the rights and freedoms of natural persons in the use of AI.[184] However, this suffered from the flaw of leaving fundamental human rights (such as privacy, data protection and the right to protection of IP) to the good conscience and subjective views of AI

177 A key definition of AI adopted in the articles in this research includes Peter Norvig and Stuart Russell (n 6) 7, Jerry Kaplan (n 8) 68.

178 Paragraph 1.2 of the Explanatory Memorandum of the AI Act, p. 4.

179 Art 68 GDPR

180 Art IV, p. 10 of this research.

181 Art 51 GDPR.

182 Paragraph 1.2 (p. 3) AI Act.

183 Articles III and IV of this research.

184 Luciano Floridi and Mariarosaria Taddeo, 'What is Data Ethics?' (Phil. Trans. R. Soc, 2016) 5 <http://doi.org/10.1098/rsta.2016.0360> accessed 19 December 2021.

deployers whose goal is primarily to turn a profit.[185] However, with the introduction and anticipated adoption of the Proposal, AI users/stakeholders will no longer be subject to the ethical magnanimity of AI deployers.

The Proposal also impacts the IP law component of this research. This can be observed from the transparency requirements of the Proposal, which ought not to disproportionately affect the protection of IP guaranteed in Article 17(2) of the EU Charter.[186] In meeting this requirement of the Proposal, only the minimum necessary information ought to be processed. Furthermore, the AI Board is also expected to respect the confidentiality of information and data and to carry out their tasks in such a manner as to protect IP rights and trade secrets.[187] Based on the above, one might infer that the EC, in drafting the Proposal, anticipates some intersection between data protection and IP law. This is because the provision considers IPR when complying with the transparency requirement and does not rule out its application to meeting the transparency requirement under the GDPR. Therefore, in meeting the (data protection) AI-related transparency requirements under the Proposal, a consideration of its impact on applicable IPR can be said to be one of the expectations of the Proposal.

185 Ibid.

186 Council of the European Union, Charter of Fundamental Rights of the European Union (2007/C 303/01), 14 December 2007, C 303/1 <https://www.refworld.org/docid/50ed4f582.html> accessed 19 December 2021.

187 Art 70 of the proposal for the AI Act.

5 EMPIRICAL ANALYSIS AND RESOLUTIONS

5.1 Resolution of Specific Concerns: An Empirical Analysis

Interviews are a subset of the legal empiricism methodology, which is widely used to elicit information relevant to the thesis statement of existing or ongoing research.[188] As a result of the multidisciplinary nature of this research, interviews have been conducted per the legal empiricism methodology. One of the reasons behind the adoption of the interview method over other possible methods with the legal empiricism methodology (for instance, the use of questionnaires) is the possibility for the interviewer to immediately probe further and/or ask follow-up questions as needed. This use of interviews ensures that the research considers the diverse views of relevant stakeholders and practitioners regarding the issues raised in this research. The interviews also provide insight into the industry expectations/prevailing practices regarding the issues discussed herein, thereby creating an avenue for comparison and deeper analysis. The interviews were semi-structured with defined but open-ended questions, allowing the interviewees to answer the questions concretely while expressing relevant opinions within the scope of the interview questions.[189] The semi-structured nature of the interviews helped keep the interview focused on its intended theme while also eliciting relevant information from the interviewees. An interview protocol containing the interview questions was used for this interview.[190] The interview protocol was divided into a general section and two main categories - the right to data portability and IP-related questions. The right to data portability served as a use case to depict how the exercise of data subject rights could result in data protection issues during data reidentification.

On the other hand, the IP-related questions were mainly used to explore the interest of interviewees (and, by extension, their organisations) in the ownership of the works generated by their AI systems. The consideration of data protection and IP law issues in two separate sections in the interview protocol was mainly used to address the research's multidisciplinary nature and prevent the issues from being muddled up. Four interviews were conducted with interviewees selected across

188 Steinar Kvale, Interviews An Introduction to Qualitative Research Interviewing (Sage Publications, 1996).

189 Jennifer Mason (n 50). See also Barbara DiCicco-Bloom and Benjamin F. Crabtree, 'The qualitative research interview' (2006) Med Educ. 40, 314–21.

190 Appendix 1 of this synthesis.

different spectra of the field of AI, including academia, AI manufacturing, and the legal services industry. Based on their experience with AI, the interviewees were able to give practical and pragmatic insights into some of the issues considered herein. In analysing the views of these experts, this research takes the need of the industry and the practical realities of AI practitioners and stakeholders into perspective. The selection of interviewees was based on their relevant experience in the field of AI. The interviews were recorded and transcribed to allow for a more detailed analysis. See Appendix II for a description of the interviewees.

The multi-disciplinary nature of this research and its use of the doctrinal research method necessitated an evaluation of its findings, especially those that arose from the analysis of various legal instruments and literature. These interviews have therefore been used to obtain the perspectives of AI experts and practitioners on the conclusions and findings in the four articles making up this research and other relevant issues (further) addressed in this synthesis. By so doing, this research mainly takes existing practical views and solutions into perspective in evaluating some of its findings and recommendations.

5.1.1 Interview analysis and discussion

To maintain the anonymity of the respondents, the relevant findings of this research have been considered in aggregation, as against singularly analysing each interview. Therefore, this analysis has been carried out in three sections per the three sections highlighted in the interview protocol.

5.1.1.1 General Questions

This section of the interview protocol highlights some questions about the interviewee's background. The section sought answers to questions such as the role and professional background of the interviewee in the development, deployment, or sale of AI; whether the interviewees' employer or company developed and/or sold AI; the type of AI in question; the kind of output generated by the AI system; the type of data (both personal and non-personal) processed by the system; whether the company's data subjects/customers were individuals or corporate entities or both. Whether data subjects are, natural or legal persons is relevant in determining the impact of data subject rights on AI. This is because only natural persons can make such requests. The answers to the questions in this section vary subject to the job description and profile of the interviewee. See Appendix III.

5.1.1.2 The Right to Data Portability

The second category of questions in the interview protocol is centred around data portability and concerns that might arise in exercising this right should IPR protect (AI-generated) data. The focus on the right to data portability was used to highlight and evaluate the intersection between IP and data protection law within

an AI context. The questions in this section identify how AI industry participants and stakeholders are resolving, handling, and/or preparing for possible occurrences of data reidentification when data subjects exercise their data subject rights. These effects of data reidentification are examined using the exercise of data subject rights, particularly the right to data portability. The right to data portability was used to highlight the possible sharing of a company's business data with competitors in the data reidentification scenario addressed in this research. Some of the issues considered in this section include whether data subjects (or customers) have transferred their data to any other organisation based on the right to data portability, the type of data involved, fears as to competitors' access to business information and how this was resolved; general expert opinion of the interviewee on these issues; etc.

The interviews revealed that the interviewees had neither received nor responded to any requests for the exercise of the right to data portability, including within the context of this research. Arguably, this could be because the right to data portability has not enjoyed the kind of attention that other data subject access rights have received. According to Interviewee I, the rationale behind this lies in the "lack of standard interfaces and technological interoperability among companies". This affirms the futuristic nature of the issues considered herein and the necessity of proactively addressing them before they manifest. Furthermore, the lack of experience in this scenario among interviewees made other questions that would ordinarily have been followed under the interview protocol inconsequential. Such questions include the nature of data transferred, the risks encountered, how they were addressed, etc. However, this research still took advantage of the expertise of the interviewees by seeking their expert opinions on the possible impact of the right to data portability in the event of data reidentification. In expressing their expert opinions, the interviewees acknowledged that the right to data portability could pose some challenges within the context of the data reidentification of IPR-protected (AI-generated) data. The interviewees identified the (remote) possibility of the business information of controllers being leaked to their competitors as one of such challenges that could arise. In proposing a solution, Interviewee I believed that "any attempt to mitigate such potential risk must take the interest of the data subject and the data controller into consideration". Interviewee I further noted that "in balancing the competing interests between the controller and the data subject, the interests of the data subject ought to be prioritised because of their weaker position in the relationship".[191]

Interviewee II believed that implementing the right to data portability might be a bit challenging because of "the lack of standard interfaces and data sets between

191 Companies would typically be protected by more regulations and IP rights (such as patents) giving them a protective edge over individuals. They would also usually have more assets and means to lobby and project their position more than individuals.

companies". Interviewee II acknowledged the potential risks that might be posed by the right to data portability, which might (even though remotely), pose risks to the confidential information/business structures of the data exporting organisation.[192] As a solution, interviewee II suggests "exporting only the data provided by the data subject without including datasets derived therefrom, especially when it constitutes a trade secret".[193] According to the interviewee, this will necessarily result in the right to data portability being treated just like the right to access due to the lack of interoperable systems. Interviewee II also noted that there had been instances where a data subject sought to send their data to their employer, even though not within the context of the right to data portability. In those instances, the data subjects were asked to download relevant apps and services and subscribe through the regular channel owing to the lack of interoperable systems that make data portability a practical impossibility. The interviewee proposed the standardisation of the interoperability of systems as per the right to data portability but doubts the value of this activity based on the almost non-existent data portability requests presently. Interviewee III was of the view that "for entities like law firms bound by a duty of confidentiality, it is more expedient for them to rely on the confidentiality requirement as a basis for refusing to enforce the right to data portability when it might result in the violation of their confidentiality obligation".

Based on the findings above, data subject requests (generally, and the right to data portability particularly) are yet to practically pose a concern upon data reidentification. However, the interviewees acknowledged the reality of such an occurrence. In the absence of practical examples, AI experts have accepted the problems identified in this research regarding the reidentification of IPR-protected (AI-generated) data and the exercise of data subject rights, particularly the right to data portability. The experts have proposed measures such as balancing competing interests and standardisations, which harmonise the data to be exported, as possible solutions for resolving the identified concerns.

5.1.1.3 IP Related Questions

The third part of the interview protocol focussed on IP-related questions to obtain the view of AI experts on the ownership of AI-generated works. Some of the questions under this section sought to clarify the necessity/desirability of protecting AI-generated data/works and companies' current approach to protecting AI and AI-generated works.

192 The data exporting organisation in this context will be the entity sending out the data on the request of the data subject.

193 Derived data is produced from other data in a relatively simple and straightforward fashion. For example, by calculating customer profitability from the number of visits to a store and items bought. The Information Commissioner's Office (n 5) 12.

Concerning the necessity of IP protection for AI-generated works, the interviewees generally opined that such protection might be necessary, though from different perspectives. For instance, Interviewee II expressed the desire to earn fees from AI-generated works as a rationale for protecting such works. Interviewee IV believed that "such recognition was necessary to ensure that the AI system (and/or its manufacturers) are acknowledged for their role in the generation of the relevant data". However, Interviewee III took a deferring opinion proposing that only AI systems that accomplish complex tasks should be protected. These opinions of these interviewees align with the position of this research that AI-generated works ought to be granted some form of IP protection for reasons including incentivising AI manufacturers. Interviewee I was of the view that "IP rights attributable to AI-generated works should be ascribed to persons who have made significant and original contributions to AI-generated works including developers of the AI system, anyone who contributed training data to the system, and its users".

The interviewees showed varying levels of interest in the ascription of IPR to the AI-generated works. However, interviewee III believed that AI should only be granted IP rights when their output is specific and not amounting to routine work. Interviewee IV opined that IPR should only be ascribed to AI to protect the moral rights of the AI (or AI manufacturers). The interviewees' views suggest that the 'public domain' argument, proposed in scholarly works and examined in Article I of this research, is not considered a solution amongst the interviewed AI industry practitioners.[194] Even though they have not expressly spoken against the 'public domain' proposition, their interest in IPR protection of AI-generated works is antithetical to AI-generated works. Despite the general interest in protecting AI-generated works, only one of the four entities affiliated with the interviewees (i.e. Interviewee II) currently actively protects its AI-generated works as of today. The rationale is for companies to earn income from their IP rights to maintain their platform and provide quality services. In other words, the company is remunerated through the royalties earned from the said IP rights. The other three are not actively protecting AI-generated works for reasons which include the difficulty of some of such works meeting the eligibility requirements for IP protection.

The last question under this section pertains to recognising AI as a separate legal person. This question is relevant because it partly determines if AI systems can be ascribed to IPR. This is because IPR must be vested in a natural/legal person, and AI systems are neither of these. This question sought to discover the expert opinion of AI industry practitioners on ascribing legal personhood to AI systems. The opposition of the interviewees' to the ascription of legal personhood to AI was justified partly because it could allow corporations to avoid liability for any misdeeds of their AI

194 The 'public domain' argument proposes that AI-generated works should be made to lie in the public domain without any possibilities of IP protection. See Tanya Aplin and Giulia Pasqualetto (n 62).

creations.[195] Interviewee I posited that AI systems are, at best autonomous and not independent. This implies that even when AI systems can autonomously produce (original) works or novel, inventive inventions, they might not be independent because of the human input needed to create AI systems and design their execution of tasks. In other words, AI systems will always be designed to execute goals outlined by others, making them capable of autonomy but not independence. Interviewees II and III were also opposed to the ascription of legal personhood to AI for IPR purposes and, in general, because of the human input needed therein, which meant that the systems were incapable of acting independently. However, the interviewees were open to the possibility of such AI systems being granted legal personhood in the futuristic case of strong AI and where a structure like the corporate legal personality was introduced. Therefore, the interviewees are more open to futuristic strong AI systems owning IPR but would not recommend the same for weak AI systems as we largely have today.

195 Such liability could arise where AI systems violate laws by, (for e.g.), exceeding the instructed scope of the purpose(s) for processing personal data.

6 The Redefinition of Personal Data in the Age of AI: A Feasible Solution?

This research has identified specific research questions that have been variously considered through the articles and in legal empiricism. These research questions contribute variously to the theme of this research which revolves around the data protection risks that could arise from the IPR protection of AI-generated data. This research has established that (personal) data is fundamental to the operation of AI and vice versa. From some of the considerations in this research, it has been further established that AI magnifies the already blurred distinctions between personal and non-personal data by further increasing the possibility of data re(identification) through various means, which include the linking of non-personal datasets in a manner that can identify data subjects. This creates a scenario that potentially limits the data categories that can be (lawfully) processed by AI systems. This is particularly important because most of the data required by AI systems (for machine learning, for instance) are non-personal data. Since these datasets are used to train AI systems, it is a customary and accepted practice that personal data is not used in a test environment.[196] In such instances, one can expect that AI developers would rather use non-personal data for machine learning as this helps avert many risks that might arise from personal data processing. The blurring of the distinction between personal and non-personal data, further amplified by AI, can result in the unintended processing of personal data when non-personal data would have been sufficient. In addition, it can potentially slow down the AI development process because AI developers might be forced to process personal data and, by extension, comply with data protection law when they intend to process non-personal data merely.

Furthermore, complying with data protection law in circumstances of this nature will potentially limit the personal data available for AI development because a justifiable legal basis might be inapplicable. For instance, not every data subject will be willing to consent to using their personal data in such cases, thereby limiting the volume of data available for processing.[197] As a means of ameliorating this

196 In accordance with ISO standards, production data should not be used in a test environment. Within the context of this research, this means that personal data ought not be used for simulations and training, as this will constitute using production data in a test environment. See International Organization for Standardization (ISO) 27001:2022, Information security, cybersecurity and privacy protection — Information security management systems — Requirements, control 8.31. <https://www.iso.org/standard/82875.html> accessed 2nd January 2022.

197 It is most likely that consent is the most feasible legal basis in this use case. See Article 6 (1) (a) – (f) GDPR, particularly paragraph (a).

possible consequence, this research proposes that it might be necessary to revisit the definition of personal data to ensure that its current (broad) definition does not hinder the development of AI. Under the current definition of personal data, any data which can directly or indirectly lead to the identification of natural persons, including data which might be capable of reidentification, would require data protection compliance. This is because any data that identifies a natural person, even remotely, amount to personal data. This approach to the definition of personal data was recognised even before the coming into force of the GDPR. According to the Article 29 Working Party, any information would ordinarily qualify as personal data except when identifying natural persons is impossible.[198] Datasets such as IP addresses from technological devices,[199] web traffic surveillance tools and geolocation services,[200] etc., are treated as personal data even though they are only linked to and do not directly identify natural persons. Therefore, the identifiability of a natural person, either directly or through the combination of datasets, is the defining criterion that qualifies a dataset as personal data.[201] This remains so even for data categories that pose no apparent risks within the context of the data processing operation. The definition of personal data is so expansive that it covers not only data that readily identify a natural person but also data that may potentially identify a natural person. It is contended that with the technological advancements, possibilities for data reidentification abound, and most datasets that AI systems will process may expressly be personal or non-personal data that can be reidentified as personal data.

To avoid the paucity of data necessary for AI systems' processing activities, this research unequivocally recommends the redefinition of personal data. This research proposes that for data categories which do not directly identify natural persons, such data categories can be considered as less critical personal data which can be processed with more lenient conditions, even possibly without compliance with some of the principles of data protection law. This research will refer to these data categories as 'secondary personal data'. Data categories such as movie reviews, customer feedback, etc., when they do not directly identify natural persons and/or are not intended for such purpose, could be treated with more data protection leniency. Hypothetically, relevant entities can be made to apply less stringent data protection principles to these processing activities.

198 Article 29 Working Party, 'Opinion 4/2007 on the Concept of Personal Data' (WP 136, 20 June 2007) 6.

199 Article 29 Working Party, 'Privacy on the Internet – An Integrated EU Approach to On-line Data Protection' (WP 37 21, November 2000), 21.

200 Article 29 Working Party, 'Opinion 13/2011 on Geolocation Services on Smart Mobile Devices' (WP 185, 16 May 2011).

201 Article 29 Working Party (WP29), 'Opinion 4/2007 on the Concept of Personal Data' (WP 136, 20 June 2007), 6.

Regarding the application of more lenient data protection principles, regulation must provide for the standardisation of which data protection principles may be dispensed and the extent to which this may be done to allow for uniformity in applying these principles. For instance, the lawfulness principle and the transparency principle could be dispensed with in the processing of 'secondary personal data' in AI systems. In such a scenario, it is necessary also to designate an ad hoc data controller who will resume data processing activities if data protection law becomes applicable because of data reidentification.[202] Consequently, this approach will limit the concerns that can arise from the processing of 'secondary personal data' while also making data available for developing AI systems.

In addition, specific technical and organisational measures to restrict data access, prohibit data repurposing and ensure data security ought to be adopted to limit the chances of data reidentification. An examination of data reidentification experiments[203] would reveal that such reidentification was possible in some cases because the relevant data sets were publicly available.[204] Therefore, it is necessary to adopt specific technical and organisational measures to keep such data from public access. By so doing, AI systems will have access to relevant (secondary personal) data while still protecting the rights of data subjects. It is argued that the legal, ethical, and contractual protection of 'secondary personal data' will be sufficient to protect personal data, especially when the data subject's identification is neither necessary nor intended by the data controller. Practically, AI developers will typically not be interested in identifying natural persons, especially when they do not have the means to identify such natural persons. It would be more time and capital-intensive for AI developers to identify data subjects except when there is a need to make such identification. In such cases where data (re)-identification is necessary for data processing, such processing would be treated as a personal data processing activity and will not be treated as 'secondary personal data'. A proposition with a similar objective has previously been advanced by Hon et al., who proposed adopting a technologically neutral and accountability-based approach to reduce the possibility of personal data identification, thereby reducing the scope of data that is considered personal data.[205] These scholars suggest two approaches to achieve this objective. The first approach involves the adoption of appropriate technical and organisational measures to reduce the possibility of the identification of natural persons. The second approach assesses the possible risk and severity of harm to data subjects. In the event

202 This recommendation has been made in Article II of this research.

203 See Article II of this research.

204 Latanya Sweeney, (n 111).

205 Kuan Hon, W, Christopher Millard, and Ian Walden, 'What is Regulated as Personal Data in Clouds?', in Christopher Millard (ed.), Cloud Computing Law (Oxford, 2013; online edn, Oxford Academic, 23 Jan. 2014), 211, 214-222, 227-228 <https://doi.org/10.1093/acprof:oso/9780199671670.003.0007> accessed 27 October 2022.

of a sufficiently compelling level of seriousness, appropriate measures proportionate to identified risks and harm must be taken.[206]

For this proposition to work, it is necessary to outline an objective framework for determining when 'secondary personal data' can identify a natural person and/or when its usage can violate the rights and freedoms of natural persons. The necessity of adopting such a uniform approach lies in the fact that what translates to a data protection violation might vary among various data subjects and stakeholders subject to their appreciation and comprehension of their data protection (related) rights.[207] This is irrespective of the clarity of the GDPR on possible activities that can amount to violations under it. An example of the subjective interpretation and appreciation of data protection rights could be seen in possible reactions to the processing of IP addresses. IP addresses have undoubtedly been recognised as personal data even before the GDPR.[208] However, it is reasonable to expect that while some individuals identify the value in data controllers upholding the protection of their IP addresses, other data subjects might not see any value in its protection, even though the GDPR guarantees its protection. However, irrespective of whether data subjects appreciate the necessity of its protection, the data protection concerns that flow from this activity remain the same.[209] Therefore, expressly outlining an objective framework for determining when 'secondary personal data' will expressly impact data protection law will reduce any risks arising from a subjective interpretation of the processing activity.

Furthermore, such an objective framework could be benchmarked against the possibilities of reidentification balanced against the technical and organisational measures which reduce the likelihood of such reidentification. If adopted, this approach will narrow the definition of personal data and make more data available in AI systems. This will prevent a situation where data protection law poses any form of hindrance to the development of AI systems.

The proposition for the regulation of 'secondary personal data' above is also relevant to the IPR aspect of this research. In Article I of this research, a proposition was made for the IPR of AI-generated works to lie in the AI system that has generated it. This research further proposed that in such an event, data subject rights might apply to such AI-generated works, particularly in the event of data (re)identification which arises from the singling out of natural persons from anonymised and/or non-personal data. Such data (re)identification might occur when such data was

206 Hon et. al. (2011), Footnote 186.

207 It is necessary to note that not all stakeholders involved in the use of the GDPR believe in its (total) benefit. See Insight Software, 'GDPR- The Good, the Bad, and the Ugly' (Insight software, 2018) <https://insightsoftware.com/blog/gdpr-the-good-the-bad-the-ugly/> accessed 29/11/2021.

208 Article 29 Working Party, (n 200). See also Case 582/14 – Patrick Breyer v Germany

209 IP addresses can be used to track them and provide geolocation services without their knowledge. See Article 29 Working Party, (n 200).

used in processing AI-generated data, particularly before the (re)identification. The proposition for the use of 'secondary personal data' can also be useful in this regard. The technical and organisational measures and the objective framework proposed above will help AI developers recognise when non-personal data can potentially single out data subjects, thereby limiting the possibility of such data being entangled with AI-generated data.

Furthermore, these technical and organisational measures will ensure that 'secondary personal data' used in processing AI-generated data are securely processed in a manner that limits data breaches that might result in the matching of non-personal datasets to single out data subjects. Therefore, AI-generated works will potentially not be subject to data subject rights. Any avenue for such a possibility will likely have been identified and averted. By extension, related arguments against AI systems owning IPR in their works because of the possibility of data subject rights will be significantly reduced.

6.1 Concluding Remarks

One clear finding from this research is that there is a need to reconsider the applicability of conventional data protection and IP law principles within AI systems. The central theme of this research is that if not adequately considered, data protection and IP law will limit the data available for use in AI systems. From a data protection law perspective, this may occur because of the broad definition of personal data. The unique framework and configuration of AI systems also requires a more tailored application of data protection law principles if data protection compliance is to be achieved. IP law, on the other hand, can pose a hindrance to the development of AI systems if such systems are not given IP rights in the works they create, thereby resulting in the disincentivisation of the manufacturers of the AI systems. One of the solutions proposed to counter the protection of AI-generated works is the contention that such works should be left to lie in the public domain. However, this research posits that such an approach might limit the patronage of human-generated works, thereby disincentivising human authors.

This research makes a big case for the concurrent regulation of data protection and IP law regarding the use of AI systems. The contention is that if not concurrently regulated, both fields of the law might result in conflicting implications within AI systems. It is not unusual for academics to argue against enacting specific rules to govern AI systems.[210] However, the inference that can be drawn from the EC is that it is necessary to have a particular set of laws regulating AI. For example, the proposal for the AI Act outlines specific rules for regulating AI systems in the

210 See Tanya Aplin and Giulia Pasqualetto (n 62).

EU. Therefore, it is necessary for there to be specific legislation for AI systems and for such legislation to address the data protection and IP concerns that have been examined in this research. From the interviews, some of the potential challenges identified under this research, such as data (re)identifiability and the attendant risks in a data subject rights context, are yet to be experienced. However, the interviewees have not ruled out the possibility of such risks. The essence of research of this nature is to drive debates that will birth legislative action on some of these futuristic but possible challenges.

Finally, it is necessary to reconsider the scope of the definition of personal data if AI is to fulfil its potential in the EU. This is because the broad spectrum of the current definition of personal data coupled with technological advancements that make data easily (re)identifiable translates into a further limitation of the datasets available for use in AI systems. Since AI systems are nothing without data, broadening the scope of data available through the redefinition of personal data will be helpful in this case. Process standardisation, the adoption of technical and organisational measures, and the use of ad hoc data controllers able to step in in the event personal data become applicable are some of the measures which can be adopted if more datasets are to be made available for use in AI systems. These measures require urgent action as society becomes progressively less capable of avoiding these concerns especially in light of the continuous proliferation of AI systems. The truth is that AI systems make data protection compliance more difficult than ever before. Despite the best efforts of all relevant stakeholders, there is no guarantee of strict privacy compliance, but the advantages of AI remains boundless. It is therefore necessary to create a legal framework that both guarantees the continuous development of AI through the availability of non-personal data, that does not harm the right to data protection. This balance is key.

7 REFERENCES

ARTICLES

1. Alan M. Turing, 'Computing Machinery and Intelligence' (1950) Mind 49 433-460.
2. Barbara DiCicco-Bloom and Benjamin F. Crabtree, 'The qualitative research interview' (2006) Med Educ. 40, 314–21.
3. Bernard Dickens and Rebecca Cook, 'Legal and Ethical Issues in Telemedicine and Robotics' (2006) 94 International Journal of Gynecology & Obstetrics 1, 73-78.
4. Brent Mittelstadt, 'Principles Alone Cannot Guarantee Ethical AI' (Nature Machine Intelligence, November 2019). <https://ssrn.com/abstract=3391293>
5. Charlotte O'brien, 'I trade, therefore I am: Legal personhood in the European Union' (2013) 50 Common Market Law Review, (6) 1643–1684.
6. Chrispas Nyombi, 'Lifting the Veil of Incorporation Under Common Law and Statute' (2014) International Journal of Law & Management Vol 56 Issue 1, 6-81.
7. Christopher Rees, 'Who Owns Our Data?' (SSRN, 2013) <https://ssrn.com/abstract=2310662>
8. Cyrill P. Rigamonti, 'Deconstructing Moral Rights' 47(2) (Harvard International Law Journal, 2006) 362-367.
9. Daniel James Greiner, The new legal empiricism and its application to access-to-justice inquiries. Daedalus 2019, 148 (1): 64–74.
10. Emmanuel Salami, 'Autonomous transport vehicles versus the principles of data protection law: is compatibility really an impossibility?' (November 2020) International Data Privacy Law, Vol 10, Issue 4, 342.
11. Herbert Zech, 'A Legal Framework for a Data Economy in the European Digital Single Market: Rights to Use Data' (2016) Journal of Intellectual Property Law & Practice, Vol. 11.
12. Jane C. Ginsburg, 'The Concept of Authorship in Comparative Copyright Law' (2003) 52 DEPAUL 1072.
13. Jon Kleinberg, Jens Ludwig, Sendhil Mullainathan, Cass Sunstein, 'Discrimination in the Age of Algorithms' (SSRN, February 5, 2019) <https://ssrn.com/abstract=3329669>
14. Latanya Sweeney, 'Matching Known Patients to Health Records in Washington State Data' (SSRN, 5 June 2013) <https://ssrn.com/abstract=2289850>
15. Lawrence B. Solum, 'Legal Personhood for Artificial Intelligences' (1992) 70 N.C. L. 1239.
16. Linda J. Lacey, Of Bread and Roses and Copyrights (1989) Duke Law Journal 1532-1596.
17. Luciano Floridi and Mariarosaria Taddeo, 'What is Data Ethics?' (Phil. Trans. R. Soc, 2016) 5 <http://doi.org/10.1098/rsta.2016.0360>
18. Luciano Floridi, 'The European Legislation on AI: a Brief Analysis of its Philosophical Approach' (2021) 34 Philos. Technol. 215–222 (2021) <https://doi.org/10.1007/s13347-021-00460-9>
19. Michael Quartermain, 'Echocardiogram' in E. Alboliras, Z. M. Hijazi, L. Lopez and D. J. Hagler (eds), Visual Guide to Neonatal Cardiology (Wiley Online Library, 2018).
20. Michael Veale, Frederik Zuiderveen Borgesius, 'Demystifying the Draft EU Artificial Intelligence Act — Analysing the good, the bad, and the unclear elements of the proposed

approach'(2021) Computer Law Review International vol. 22, no. 4, 97-112.
21. Pamela Samuelson, 'AI Authorship?' (Communications of the ACM, 2020) Vol 63, Number 7, 22.
22. Péter Mezei, 'From Leonardo to the Next Rembrandt – The Need for AI-Pessimism in the Age of Algorithms' (SSRN, 4 May 2020). < https://ssrn.com/abstract=3592187>
23. Rim Razzouk and Valerie Shute, 'What Is Design Thinking and Why Is It Important?' (2012) 82(3) Review of Educational Research.
24. Sandra Erdelez and Sheila O'Hare, 'Legal informatics: application of information technology in law' (1997) 32 Annual Review of Information Science and Technology 367-402.
25. Steinar Kvale, Interviews An Introduction to Qualitative Research Interviewing (Sage Publications, 1996).
26. Suzanne Egan, 'The Doctrinal Approach in International Human Rights Scholarship' (UCD Working Papers in Law, 29 November 2017) Criminology & Socio-Legal Studies Research Paper No. 19/17.
27. Terry Hutchinson, 'The Doctrinal Method: Incorporating Interdisciplinary Methods in Reforming the Law' (2015) Erasmus Law Review, 8, 3.
28. Terry Hutchinson, 'Valé Bunny Watson? Law Librarians, Law Libraries and Legal Research in the Post-Internet Era' (2014) Law Library Journal (106(4) 584.
29. The principle of big purpose limitation and big data' Marcelo Corrales et al. (eds.), New Technology, big data and the law (Springer, 2017) 20.
30. Thomas Hoeren, 'Big Data and the Ownership in Data: Recent Developments in Europe' (2014) 36 EIPR 751.
31. Ziad Obermeyer, Brian Powers, Christine Vogeli and Sendhil Mullainathan, 'Dissecting racial bias in an algorithm used to manage the health of populations' 366 (2019) Science 447-453.

TEXTBOOKS

1. Ali Madani, Ramy Arnaout, Mohammad Mofrad and Rima Arnaout, 'Fast and Accurate View Classification of Echocardiograms Using Deep Learning' (Digital Science 1, 2018) 6.
2. Andreas Herrmann, Walter Brenner and Rupert Stadler, Autonomous Driving: How the Driverless Revolution will Change the World (Emerald Group Publishing 2018).
3. Chris Reed, 'Taking Sides on Technology Neutrality' (Scripted, 2007) Volume 4, Issue 3, September 2007
4. Christopher Kuner, Lee A. Bygrave, Christopher Docksey (eds), The EU General Data Protection Regulation (GDPR): A Commentary (2020).
5. Christopher Millard (ed.), Cloud Computing Law (Oxford, 2013; online edn, Oxford Academic, 23 Jan. 2014).
6. COE/FRA, Handbook on European Data Protection Law, (Publications office of the European Union, 2018).
7. Intellectual Property, Edward N. Zalta (ed), (Stanford Encyclopedia Of Philosophy, 2011).
8. Jennifer Mason, 'Linking qualitative and quantitative data analysis' in Alan Bryman, Robert G. Burgess (eds), Analysing qualitative data (Routledge, 1994).
9. Jerry Kaplan, Artificial Intelligence what everyone needs to know (Oxford publishers, 2016).
10. Jessica Berg, 'Of Elephants and Embryos: A Proposed Framework for Legal Personhood' (2007) 59 Hastings L.J. 369.
11. Lionel Bently and Brad Sherman, Intellectual Property Law (4th Edn, OUP 2014).
12. Lothar Terfloth, Simon Spycher, Johann Gasteiger, 'Drug Discovery: An Overview' in Thomas Engel and Johann Gasteiger (eds), Applied Chemoinformatics (Wiley, 2018).

13. Mandy Burton, 'Doing Empirical research, Exploring the decision making of Magistrates and Juries' in Dawn Watkins and Mandy Burton (eds), Research Methods in Law (Routledge, 2nd edn, 2018).
14. Markus Maurer, J. Christian Gerdes, Barbara Lenz and Hermann Winner (eds), Autonomous Driving: Technical, Legal and Social aspects (Springer, 2015).
15. Martin Bulmer, 'The Ethics of Social Research' in Nigel Gilbert (ed), Researching Social Life, (London: Sage, 3rd edn, 2008).
16. Martyn Hammersley and Paul Atkinson, Ethnography: Principles in Practice (Routledge, 3rd edn, 1995).
17. Mayson, French and Ryan, Company law (OUP, 2016).
18. Mor Bakhoum, Beatriz Conde Gallego, Mark-Oliver Mackenrodt, Gintarė Surblytė-Namavičienė (eds), Personal Data in Competition, Consumer Protection and Intellectual Property Law Towards a Holistic Approach? (Springer, 2018).
19. Nigel Gilbert, 'Research, Theory and Method' in Nigel Gilbert (ed), Researching Social Life (Sage, 3rd edn, 2008).
20. Paul Chynoweth, 'Legal Research' in Andrew Knight and Les Ruddock (eds), Advanced Research Methods in the Built Environment (Wiley-Blackwell, 2008).
21. Pauline Westerman, 'Open or Autonomous: The Debate on Legal Methodology as a Reflection of the Debate on Law' in Mark Van Hoecke (ed), Methodologies of Legal Research: Which Kind of Method for What Kind of Discipline? (Hart Publishing Ltd, 2011).
22. Peter Seipel, Computing Law. Perspectives on a New Legal Discipline, (Liberfolag, 1977).
23. Reva Berman Brown, 'Doing your dissertation in business and management' (SAGE Publications, 2006) <https://dx.doi.org/10.4135/9781849209069>
24. Rosa Ballardini, Kan He, and Teemu Roos, 'AI_Generated Content: Authorship and Inventorship in the age of Artificial Intelligence' <https://www.cs.helsinki.fi/u/ttonteri/pub/aicontent2018.pdf>
25. Rosa Ballardini, Olli Pitkänen and Petri Kuoppamäki (eds), Regulating Industrial Internet through IPR, Data Protection and Competition Law (Kluwer, 2019).
26. Rossana Ducato and Alain Strowel (eds), Legal Design Perspectives: Theoretical and Practical Insights from the Field (Ledizioni, 1st edn, 2021).
27. Steven Finlay, Artificial Intelligence and Machine Learning for Business: A No-Nonsense Guide to Data Driven Technologies (Relativistic, 3rd edn, 2018) 10.
28. Stuart J. Russell and Peter Norvig, Artificial Intelligence: A Modern Approach (Pearson, 3rd edn, 2010.
29. Visa A. J. Kurki, A Theory of Legal Personhood (OUP, 2019).

CASES

1. Bonnier Audio AB, Earbooks AB, Norstedts Förlagsgrupp AB, Piratförlaget AB, Storyside AB v. Perfect Communication Sweden AB, 19 April 2012.
2. C-275/06, Productores de Música de España (Promusicae) v. Telefónica de España SAU [GC], 29 January 2008.
3. C-461/10, Bonnier Audio AB, Earbooks AB, Norstedts Förlagsgrupp AB, Piratförlaget AB, Storyside AB v. Perfect Communication Sweden AB, 19 April 2012.
4. Case 582/14 – Patrick Breyer v Germany.
5. Case C-145/10 Eva-Maria Painer v Standard Verlages GmbH et al. [2013] ECLI:EU:C:2013:138, 42.
6. Case C-145/10 Eva-Maria Painer v Standard Verlages GmbH et al. [2013] ECLI:EU:

C:2013:138, 42.
7. Case C-145/10 Eva-Maria Painer v. Standard Verlages GmbH etal. [2013] ECLI:EU:C:2013:138. 42.
8. Case C-604/10 Football Dataco Ltd etal., v Yahoo! Etal [2012]ECLI:EU:C:2012:115.
9. EDÖB v Google BGE 138 II 346.
10. Football Association Premier League Ltd and Others v QC Leisure and Others (C-403/08) and Karen Murphy v Media Protection Services Ltd (C-429/08).
11. Tele2 Sverige AB v Post- och Telestyrelsen (C-203/15) EU:C:2016:970; [2017] Q.B. 771, Para 90, 102, 103, 108-110.

LEGAL INSTRUMENTS

1. Agreement on Trade-Related Aspects of Intellectual Property Rights, Apr. 15, 1994.
2. Berne Convention for the Protection of Literary and Artistic Works, September 9, 1886, as revised at Stockholm on July 14, 1967, 828 U.N.T.S. 221.
3. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions - Building Trust in Human Centric Artificial Intelligence (COM(2019)168).
4. Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data (ETS No. 108), as it appears in document CM(2018)2-final, and, as an instrument related to the Protocol, endorsed the Explanatory Report, as it appears in document CM(2018)2-addfinal;
5. Council of the European Union, Charter of Fundamental Rights of the European Union (2007/C 303/01), 14 December 2007, C 303/1 <https://www.refworld.org/docid/50ed4f582.html>
6. Directive (EU) 2019/790 of the European Parliament and of the Council of 17 April 2019 on Copyright and Related Rights in the Digital Single Market and amending Directives 96/9/EC and 2001/29/EC.
7. Directive 2001/29/EC of the European Parliament and of the Council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society.
8. Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications) OJ L 201, 31.7.2002.
9. Directive 2009/24/EC of the European Parliament and of the council of 23 April 2009 on the legal protection of computer programs (2009) OJ L111.
10. Directive 96/9/EC of the European Parliament and of the council of 11 March 1996 on the legal protection of databases [1996] OJ L 77/20.
11. European Convention on Human Rights, 1950, as amended by Protocols Nos. 11, 14 and 15 supplemented by Protocols Nos. 1, 4, 6, 7, 12, 13 and 16.
12. Marrakesh Agreement Establishing the World Trade Organisation, Annex 1C, 1869 U.N.T.S. 3; 33 I.L.M. 1197 (1994).
13. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation, GDPR) (OJ L 119/2016).
14. 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.

15. UK Copyright, Designs and Patents Act 1988 (UK CPDA)
16. White Paper On Artificial Intelligence—A European approach to excellence and trust COM(2020) 65 final <https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf>
17. WIPO Copyright Treaty, Dec. 20, 1996, S. Treaty Doc. No. 105-17 (1997); 2186 U.N.T.S. 121; 36 I.L.M. 65 (1997).

ONLINE PUBLICATIONS

1. Abigail Ng, 'Completely driverless cars are being tested in China for the first time' (CNBC, 2020) <https://cnb.cx/37CgG9J> accessed 22/07/2021. Germany is also planning to partly roll out AV in 2022.
2. Ahmed Elgammal, Meet AICAN, a machine that operates as an autonomous artist' (The Conversation, October 2018) <https://theconversation.com/meet-aican-a-machine-that-operates-as-an-autonomous-artist-104381>
3. AI HLEG, 'Ethics Guidelines for Trustworthy AI' (EU Commission website, 2019) < https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai >
4. Andres Guadamuz, 'Artificial intelligence and copyright' (WIPO magazine, October 2017) https://www.wipo.int/wipo_magazine/en/2017/05/article_0003.html
5. CBInsights, '33 Industries other than Auto that Driverless Cars could turn upside down' (CBInsights, 2018) <https://www.cbinsights.com/research/13-industries-disrupted-driverless-cars/>
6. CBInsights, '40+ Corporations Working on Autonomous Vehicles' (CBInsights, 2020) <https://www.cbinsights.com/research/autonomous-driverless-vehicles-corporations-list/>
7. Derek Jansen and Kerryn Warren, 'What (Exactly) Is Research Methodology? A Plain-Language Explanation & Definition (With Examples)' (Gradcoach, June 2020) <https://gradcoach.com/what-is-research-methodology/>
8. Human Rights Watch, 'How the EU's Flawed Artificial Intelligence Regulation Endangers the Social Safety Net: Questions and Answers' (Human Rights Watch, November 2021). <https://www.hrw.org/news/2021/11/10/how-eus-flawed-artificial-intelligence-regulation-endangers-social-safety-net>
9. Insight Software, 'GDPR- The Good, the Bad, and the Ugly' (Insight software, 2018) <https://insightsoftware.com/blog/gdpr-the-good-the-bad-the-ugly/>
10. Jonathan Flowers, 'Strong and Weak AI: Deweyan Considerations' (2019) AAAI Spring Symposium: Towards Conscious AI Systems <https://api.semanticscholar.org/CorpusID:57663042>
11. Jonathan Zittrain, 'The Hidden Costs of Automated Thinking' (The New Yorker, 23 July 2019). <https://www.newyorker.com/tech/annals-of-technology/the-hidden-costs-of-automated-thinking>
12. Jonathan Zittrain, 'What the Publisher Can Teach the Patient: Intellectual Property and Privacy in an Era of Trusted Privication' (SSRN, 2000). <https://ssrn.com/abstract=214468>
13. Karl Bode, 'Harvard Students Again Show Anonymised' Data Isn't Really Anonymous' (Techdirt, 20 February 2020) <https://www.techdirt.com/articles/20200203/07405543847/harvard-students-again-show-anonymised-data-isnt-really-anonymous.shtml>
14. Lauren Goode, 'AI made a movie and the results are horrifyingly amazing' (Wired, 2018) <https://www.wired.com/story/ai-filmmaker-zone-out/>
15. Mahesh Langa, 'Ahmedabad Doctor Performs Telerobotic Surgery on Patient 32 km Away' (The Hindu, December 2018) <https://www.thehindu.com/news/national/other-states/

ahmedabad-doctor-performs-telerobotic-surgery-on-patient-32-km-away/article25675166.ece>
16. Margaret Hagan, 'Law by design' (2016) <https://lawbydesign.co/legal-design/>
17. Mogens Blanke, Michael Henriques, Jakob Bang, 'A pre-analysis on autonomous ships' (Technical University of Denmark, 2016) 1 < https://www.semanticscholar.org/paper/A-pre-analysis-on-autonomous-ships-Summary/4eabcca691a52956f697f560dca3c1ce942781d8>
18. Rebecca Bellan, 'Germany gives greenlight to driverless vehicles on public roads' (Techcrunch, 2021) <https://techcrunch.com/2021/05/24/germany-gives-greenlight-to-driverless-vehicles-on-public-roads/>
19. Research Guides: Legal Dissertation: Research and Writing Guide (libguides.com). https://law.indiana.libguides.com/dissertationguide#:~:text=Doctrinal,%2C%20statutes%2C%20or%20regulations
20. Rolls-Royce, 'Ship Intelligence for Cargo Vessels' (Youtube, December 2014). <https:andandwww.youtube.comandwatch?v=_nApv-C7qSg&list=PLk-17K0buHIvy68TGjnSUppTq-Gi91lT->
21. Society of Automotive Engineers, 'SAE Levels of Driving Automation Refined for Clarity and International Audience' (SAE website, 3 May 2021) <https://www.sae.org/blog/sae-j3016-update>
22. The Information Commissioner's Office, Big data, artificial intelligence, machine learning and data protection, 20170904, version 2.2, 12 <https:andandico.org.ukandmediaandfor-organisationsanddocumentsand2013559andbig-data-ai-ml-and-data-protection.pdf>
23. The Norwegian Forum for Autonomous Ships (NFAS), 'Definitions for Autonomous Merchant Ships' in Ornulf Jan Rodseth and Hävard Nordahl (eds), (NFAS website, 2017) 7, 16-18. <https://nfas.autonomous-ship.org/wp-content/uploads/2020/09/autonom-defs.pdf>

REGULATORY INSTRUMENTS

1. Article 29 Working Party (WP29), 'Opinion 4/2007 on the Concept of Personal Data' (WP 136, 20 June 2007), 6.
2. Article 29 Working Party, 'Opinion 13/2011 on Geolocation Services on Smart Mobile Devices' (WP 185, 16 May 2011).
3. Article 29 Working Party, 'Opinion 4/2007 on the Concept of Personal Data' (WP 136, 20 June 2007) 6.
4. Article 29 Working Party, 'Privacy on the Internet – An Integrated EU Approach to On-line Data Protection' (WP 37 21, November 2000), 21.
5. International Organization for Standardization (ISO) 27001:2022, Information security, cybersecurity and privacy protection — Information security management systems — Requirements.

8 Appendices

I. Interview Protocol

Interview Protocol
Data Protection and Intellectual Property Issues in the use of AI

Theme and purpose of the interview: This interview seeks to enquire about selected data protection and intellectual property rights in the use of AI. The findings from this interview will be used as part of an ongoing research on data protection and intellectual property law pertaining to AI. The findings from the research will be shared with all interviewees once it is completed. The interview does not seek to identify legal and/or natural persons neither will there be any disclosure of names either of any legal and/or natural persons.

Emmanuel Salami, LLM.
PhD candidate, University of Lapland, Finland.
https://lacris.ulapland.fi/en/persons/emmanuel-salami(dc5c748e-5200-450b-84f2-01251880583a).html
Tel. +49 1624162212.
E-mail: Esalami@ulapland.fi

Please find below the themes I hope to discuss during the interview. The following questions are examples of issues I am interested in. I will not be requesting any confidential information.

P.S. Persons not directly affiliated to a company can express their professional opinion on company-specific questions. The questions will be adapted accordingly if the interviewee is not directly affiliated to a company.

General questions

- Does your company develop (and sell) AI? Or have you been involved in such process? Please elaborate.
- What kind of AI does your company develop?
- Please describe the output generated by your AI system.
- What types of (personal) data does your AI process?
- Are your customers individuals or companies or both? Please elaborate.

The right to data portability

- Have your customers transferred their data to any other organisation on the basis of the right of data portability? P:S. data portability is the right of customers to have their personal data transferred from one service provider to another. (Art. 20 GDPR).
- What type of data was involved?
- Were there any risks that your business information or processes may be shared with the other controller?
- Generally, do you or your company fear that your business information or processes may be shared with the other controller?
- What are the possible measures in place or in contemplation to prevent the sharing of your business information or processes with the other controllers?
- Have data subjects had their data transferred to your company pursuant to the right to data portability (described above)?
- Did you gain any business insights into the activities/ business information or processes of the other data controller?
- Based on your expertise, how would you rather advise policy makers to approach the regulation of the right to data portability in the use of AI.

Intellectual property-related questions

- Do you think IP right protection is necessary for the output generated from AI?
- Are you interested in retaining IP in the output of (your) AI systems. Is this necessary?
- Has your company left the output of your AI without IP (copyright) regulation? What's the reason for your company's decision?
- By retaining IP rights in the output of AI, do you or your company fear that it would be liable for any wrongs caused by the AI?
- Would your company be willing to take any liability for the wrong doing of AI you have developed?
- Do you think that AI should be recognised as a separate legal person from its developers? Please state the reason for your position?

II. Background of the Interviewees

Interview Number/ Interview Date	Employer Description	Respondents' Position/s in the Organisation
Interview I/December 8 2020	Computer science department of a top University that is actively involved in various AI education and development.	Professor of AI.
Interview II/January 7, 2021	A tech-based company that uses AI to provide some of its services.	Global data protection officer in a company that deploys AI.
Interview III/January 22, 2021	A leading law firm that advises clients on AI and has also deployed AI in its client advisory process.	Head of IT law (including AI).
Interview IV/22 February 2021	University that is focused on AI education and development.	AI academic and AI Engineer.

III. Profile of the Interviewees

Interview Number	Area of Expertise
Interview I	Computer scientist, AI and machine learning researcher. Interviewee I has been involved in various AI and machine learning projects covering computational data, sensitive personal data, location data for mobile devices, etc. Some of these AI systems have been relevant for both natural persons (data subjects) and legal persons.
Interview II	Data protection expert with a focus on technology/AI compliance. Interviewee II has been involved in various projects covering AI to edit maps, scanning datasets for analytical purposes, street-level imagery etc. Some of these AI-related projects have been relevant for both natural persons (data subjects) and legal persons.
Interview III	A lawyer with experience in AI advisory and deployment of AI for client advisory services. The relevant AI system applies to both natural and legal persons.
Interview IV	AI professor and co-founder of a data analytics company. Interviewee IV has been involved in various AI engineering research projects. Interviewee IV also uses machine learning to provide data analytics services to legal persons.