Universidade do Minho
Escola de Direito

Tiago Sérgio Cabral

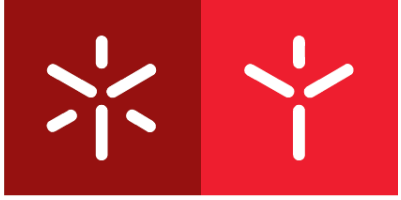
AI Regulation in the European Union: Democratic Trends, Current Instruments and Future Initiatives

outubro de 2019

AI Regulation in the European Union: Democratic Trends, Current Instruments and Future Initiatives

Tiago Sérgio Cabral

Uminho | 2019



Universidade do Minho
Escola de Direito

Tiago Sérgio Cabral

AI Regulation in the European Union: Democratic Trends, Current Instruments and Future Initiatives

Dissertação de Mestrado
Mestrado em Direito da União Europeia

Trabalho efetuado sob a orientação da
Professora Doutora Alessandra Silveira

outubro de 2019

DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS

Acknowledgements:

A few words of gratitude must be directed towards my supervisor, Professor Alessandra Silveira, for her guidance absent which it is quite unlikely that this work would have ever been completed, and for all the support provided since the far-distant times of my LL.B.

To my father for the encouragement and values passed on, and also for the phone calls reminding me that I should hurry up in finishing this work. To my mother for always quietly being there when needed, including for drawing my attention to the fact that humans need sustenance in the form of food if they desire to remain alive. Something that I frequently forgot while writing. To my sister, mostly for the patience.

To my uncle, to my aunt and to my little cousin for the warm welcome in Brussels and also for the delicious cake. To the European Commission Library's staff for always being friendly, even when I was monopolizing their printers.

To everyone who read this work before it was submitted and offered opinions and suggestions, it was undoubtedly made richer because of you. To all the friends who did not stop talking to me when, in the weeks before the deadline, the present work became (almost) the sole topic of conversation I could muster. I promise to be more interesting in the future.

In the end, it is not possible to really be fair in writing the acknowledgements section. Someone is always overlooked, even though their contribution, even if so small that it was forgotten might have been vital. Therefore, and since I am sure that I also suffer from this shortcoming, if you in any way contributed or inspired this work, please have my sincerest appreciation.

Any mistake is solely due to the Author's limitations and cannot be attributed to any of the abovementioned.

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

AI Regulation in the European Union: Democratic Trends, Current Instruments and Future Initiatives

ABSTRACT:

It is expected that the development of AI will create plenty of opportunities for economic and social progress. However, with it will also come numerous challenges. To properly reap the benefits of this world changing technology new legal concepts, principles, rules and interpretations are needed. The European Union, in particular, is aiming for a position as a leader in AI development and the world's standard-setter for future AI Regulation. Therefore, studying this question from the European perspective is of extreme practical and academic relevant.

This Thesis is divided in two parts. Part I is entitled "Efforts for AI and Machine Learning Regulation" and is outward and future focused. In Part I we will study what is AI and assess how to the European Union can regulate it in a democratic and acceptable manner. Furthermore, we will examine the current trends for AI Regulation in the EU, through the documentation released by the European Institutions and critically analyse the positions herein. Measures by Member States and trends related to AI Regulation will also be analysed along with the ones coming from the most relevant global competitors.

Part II is entitled "Current EU Legal Framework and Challenges" and it is inward and present focused. In Part II we will examine the legal instruments already available at the European level that regulate (certain areas of AI). Our focus will be on the General Data Protection Regulation, the new Consumer Protection Directives (Directive 2019/770/EU and Directive 2019/771/EU) and the Product Liability Directive.

At the end of this Thesis we will present some conclusions and policy suggestions that we believe may be of assistance in ensuring an adequate, effective and proper future AI Regulation in the European Union.

AI Regulation in the European Union: Democratic Trends, Current Instruments and Future Initiatives

RESUMO:

É expectável que dos crescentes desenvolvimentos na área da IA advenham abundantes oportunidades de progresso económico e social. Contudo, é certo que daqui também resultarão numerosos desafios. Por forma a adequadamente colher os frutos desta tecnologia, novos conceitos, princípios, regras e interpretações são necessárias na área do Direito. Em particular, a União Europeia, procura posicionar-se como líder no desenvolvimento de IA e como *standard-setter* no âmbito da criação de novas regras regulatórias nesta área.

Este Trabalho encontra-se dividido em duas partes- A Parte I é intitulada de *"Efforts for AI and Machine Learning Regulation"* e foca-se no exterior e no futuro. Nesta Parte, analisaremos o que é a IA e estudaremos de que maneira poderá a União Europeia estabelecer regras democráticas e aceitáveis para a sua regulamentação. Ademais, analisaremos as tendências regulatórias relativas a esta tecnologia na União Europeia. Medidas implementadas por Estados-Membro e tendências regulatórias nacionais serão também estudadas, assim como as advindas de países terceiros que sejam particularmente relevantes.

A Parte II é intitulada *"Current EU Legal Framework and Challenges"* e foca-se no interior e no presente. Na Parte II examinaremos os instrumentos legais que existem, no presente, e que regulam (pelo menos parcialmente) aspectos relevantes para a IA. O nosso estudo focar-se-á no Regulamento Geral sobre a Protecção de Dados, nas novas Directivas para Defesa do Consumidor (Directiva 2019/770/UE e Directiva 2019/771/UE) e na Directiva relativa à Responsabilidade do Produtor.

No final deste Trabalho, apresentaremos algumas conclusões e sugestões que, acreditamos, poderão ser pertinentes na consagração de um regime adequado e eficaz para a regulamentação da IA na União Europeia.

List of Most Frequent Abbreviations and Acronyms ..... **x**
Introduction **1**

PART I – EFFORTS FOR AI AND MACHINE LEARNING REGULATION **5**

Chapter I – Preliminary Concepts **6**
§ 1. Artificial Intelligence 6
§ 2. Narrow AI vs General AI 12
§ 3. Software Agents 13
§ 4. Machine Learning and Knowledge Engineering 16
§ 5. Schools of Machine Learning 18
§ 5.1. Symbolic AI 18
§ 5.2. Evolutionaries 19
§ 5.3. Connectionists 20
§ 5.4. Analogizers 21
§ 5.5. Bayesians 22
§ 6. Supervised, Unsupervised, Semi-Supervised and Reinforced Learning 24
§ 6.1. Supervised Learning 24
§ 6.2. Unsupervised Learning 26
§ 6.3. Semi-Supervised Learning 27
§ 6.4. Reinforcement Learning 28
§ 7. Black Box Algorithms 29

Chapter II – Legislative procedure in the European Union **31**
§ 1. Ordinary Legislative Procedure 32
§ 1.1. Trilogues 37
§ 2. Special Legislative Procedures 44
§ 2.1. Consent Procedure 44
§ 2.2. Consultation Procedure 47
§ 3. Exceptions 50
§ 4. *Passerelle* Clauses 51
§ 5. Citizens' Initiative 52

Chapter III – Current Trends for AI in the European Union **54**
§ 1. Preliminary Work 54
§ 1.1. European Council Conclusions – 19 October 2017 54
§ 1.2. The European Parliament's Legal Affairs Committee European Civil Law Rules in Robotics Study 55
§ 1.3. The European Parliament's Civil Law on Robotics Resolution 58
§ 1.3.1. General Principles Concerning the Development of Robotics and Artificial Intelligence for Civil Use 59
§ 1.3.2. Research and Innovation 62
§ 1.3.3. Ethical Principles 62
§ 1.3.4. Creation of a European Agency 63
§ 1.3.5. Intellectual Property Rights and the Flow of Data 64
§ 1.3.6. Standardisation, Safety and Security 65
§ 1.3.7. Autonomous Means of Transport 65
§ 1.3.7.1. Autonomous Vehicles 65
§ 1.3.7.2. Drones 67
§ 1.3.8. Care Robots 68
§ 1.3.9. Medical Robots 69
§ 1.3.10. Human Repair and Enhancement 69

§ 1.3.11. Education and Employment....70
§ 1.3.12. Environmental Impact....71
§ 1.3.13. Liability....71
§ 1.3.14. International Aspects....73
§ 2. The Commission Starts Building the Foundation for a Specific Legal Framework....74
§ 2.1. Communication from the European Commission: Artificial Intelligence for Europe....74
§ 2.1.1. Commission Staff Working Document: Liability for Emerging Digital Technologies....80
§ 2.2. Communication from the European Commission: Coordinated Plan on Artificial Intelligence....82
§ 2.2.1. Cooperation between the EU, Member States and Stakeholders....83
§ 2.2.2. Regulatory Initiatives....84
§ 2.2.3. Sectorial Concerns....85
§ 2.3. Ethics Guidelines for Trustworthy AI....86
§ 2.3.1. Communication from the European Commission: Building Trust in Human-Centric Artificial Intelligence....90
§ 2.4. Definition of AI HLG....90
§ 2.5. Policy and Investment Recommendations for Trustworthy AI....91
§ 2.5.1. Establishing an Appropriate Governance and Regulatory Framework....95

Chapter IV – Development and AI-related initiatives in Non-EU Countries....99
§ 1. United States of America....99
§ 2. China....103
§ 3. Japan....110

Chapter V – Development and AI-related initiatives in EU Member States....112
§ 1. France....112
§ 1.1. The Villani Report (For a Meaningful Artificial Intelligence: Towards a French and European Strategy....113
§ 2. Germany....116
§ 3. United Kingdom....119
§ 4. Italy....120
§ 5. Spain....121
§ 6. Portugal....122
§ 6.1. The Portuguese AI Strategy....123
§ 7. Finland....125
§ 8. Estonia....126
§ 9. The Netherlands....127
§ 10. Sweden....128
§ 11. Belgium....129

PART II – CURRENT EU LEGAL FRAMEWORK AND CHALLENGES....130

Chapter I – Greedy computers: machine learning and data processing....131
§ 1. The GDPR....131
§ 1.1. A Quick Look into the Relationship between AI and the General Principles of Data Protection....134
§ 1.2. Who is the Data Controller?....135
§ 1.3. The Right to Erasure (also known as the Right to be Forgotten)....137
§ 1.3.1. Under Directive Directive 95/46/EC and the Google Spain Judgment....137
§ 1.3.2. Under the GDPR....140
§ 1.3.3. The Right to Erasure and AI....144
§ 1.3.3.1. Deleting Data in Computers....144

§ 1.3.3.2. Recovering Deleted Files 145
§ 1.3.3.3. Inputting Data and Deleting Data in AI 146
§ 1.4. The Right to Rectification 153
§ 1.5. The Right to Object 154
§ 1.6. The Right to Explanation 155
§ 1.6.1. Legal Basis under the GDPR 155
§ 1.6.2. Preliminary Question: What is a Based Solely on Automated Processing? 156
§ 1.6.3. Preliminary Question: What is a Decision that Produces Legal Effects Concerning him or her or Similarly Significantly Affects him or her 158
§ 1.6.4. Preliminary Question: Recitals Under EU Law 159
§ 1.6.5. Preliminary Question: Algorithmic Bias and why the Right to Explanation is a Key Tool to Fight it – 4 Stories 162
§ 1.6.5.1. The Misogynistic AI 162
§ 1.6.5.2. Anti-Semitic AI 164
§ 1.6.5.3. The Racist AI: Part I – On Your Social Network 165
§ 1.6.5.4. The Racist AI: Part II – Imprisoning You 167
§ 1.6.6. The Right to Explanation under the GDPR: Rights to Information and Access 169
§ 1.6.6.1. Under the Rights to Information and Access: Article 13 - Limitations 170
§ 1.6.6.2. Under the Rights to Information and Access: Article 14 - Limitations 172
§ 1.6.6.3. Briefly: Valid Consent 174
§ 1.6.6.4. Preliminary Question: Should Information Given Under Article 15 Be the Same as Under Articles 13 and 14? 175
§ 1.6.6.5. Under the Rights to Information and Access: Article 15 177
§ 1.6.7. The Right to Explanation under the GDPR: Automated Individual Decision-Making 179
§ 1.7. Automated Decision-Making and the Right to Obtain Human Intervention 183
§ 2. Shortly: The Charter of Fundamental Rights of the European Union 185

Chapter II – AI-enabled Devices and Services and Consumer's Rights: A Short Introduction to the EU's new Legal Framework 186
§ 1. The Sale of Goods Directive 186
§ 2. The Digital Services Directive 188

Chapter III – Product Liability 192
§ 1. The Product Liability Directive 192
§ 1.1. Current Framework 192
§ 1.2. Potential Shortcomings on AI Regulation 196
§ 1.2.1. Is AI a Product or Service? 196
§ 1.2.2. Defences 200
§ 1.2.2.1. The State-of-the-Art Defence 200
§ 1.2.2.2. The Later-Defect Defence 203
§ 1.2.3. Proving Causality 204
§ 1.2.4. Limited Damages 206
§ 1.3. The Future Commission Guidance on the Product Liability Directive 207
§ 2. Alternative Regulatory Approaches 208
§ 2.1. Amending the Product Liability Directive 208
§ 2.2. Mandatory Insurance Schemes and Compensation Funds 209
§ 2.3. Why Do We Think That a Specific Legal Status is not the Solution (for now)? 211
Conclusions and Policy Suggestions 215

References 221
1. Articles 221
2. Books 235

3. Government Publications....242
4. Case-Law 250
5. Online Resources....253

List of Most Frequent Abbreviations and Acronyms

AEPD	Agencia Española de Protección de Datos
AI	Artificial Intelligence
Charter	Charter of Fundamental Rights of the European Union
CNIL	Commission nationale de l'informatique et des libertés
CNPD	Comissão Nacional de Protecção de Dados
Council	Council of the European Union
DPC	Data Protection Commission
DPIA	Data Protection Impact Assessment
EC	European Commission
ECJ	(European) Court of Justice
ECON	European Council
EDPB	European Data Protection Board
EGC	(European) General Court
EP	European Parliament
EU or Union	European Union
GDPR	General Data Protection Regulation
HLG	High-Level Expert Group on Artificial Intelligence
ICO	Information Commissioner's Office
ML	Machine Learning
OLP	Ordinary Legislative Procedure
SCS	Social Credit System
TEU	Treaty on European Union
TFEU	Treaty on the Functioning of the European Union
WP29	Article 29 Data Protection Working Party
XAI	Explainable AI

Introduction

8:37 AM, your alarm clock rings. By listening to you while you slept and checking whether you are in deep or light sleep it knows that it is the perfect time to wake you up and selects a song from your Spotify library to do so. The choice: "It's Beginning to Look a Lot Like Christmas", by Michael Bublé. Since your alarm is connected to your Amazon account, it knows that you were shopping for Christmas presents the day before – even though it is still October your love for Christmas cannot be stopped by such trivialities – and, as such, it will certainly fit your spirit! You try to get just five more minutes in bed, but your alarm kindly reminds you that you have a meeting at 10:00 AM and, in average, you take 45 minutes getting dressed in the morning, 30 communing to work and still like to get an early chat with your co-workers before properly diving in[1].

You give in, get out of bed and go straight to your morning routine. Luckily by connecting to the sensors on your smart watch, your shower is balanced perfectly between hot and cold and you leave feeling refreshed. Additional good news… your AI-powered toilet analysed your body waste and it appears that there is no signal of kidney or bladder issues, infections or cancer[2].

A notification appears on your phone: it is your toaster telling you that your toast is ready for consumption, and suggesting your favourite brand of orange juice to go along. You decide to follow the suggestion, take the juice out of your fridge and notice that it is already half empty (or maybe half full)[3]. No reason to despair though, immediately after you put it back, your fridge connects online and places an order for more. It checks your agenda and finds that your nephews are coming for Halloween and they also enjoy the same brand as you – you always buy extra when they come after all – so it endeavours to ensure that your stock will be full by that time .

You finish your morning chores and leave home. An autonomous car will drive you to work, while you start preparing your meeting in the back[4]. Nostalgia hits you as you

---

[1] Intelligent alarms are nothing new, *Bonjour* a French project was particularly innovative when it appeared (it was shut down in 2019), but Google Home, Amazon Alexa (Echo), similar devices or even your mobile phone can work as intelligent or AI-enabled alarm clocks. *See*, "Bonjour smart alarm clock mysteriously killed off, despite $1m crowdfunding success", Alistair Charlton, accessed September 10, 2019, https://www.gearbrain.com/bonjour-smart-alarm-clock-killed-2626720300.html.

[2] Absolutely possible. *See*, "Artificial Intelligence in Your Toilet. Yes, Really!", "Bernard Marr", accessed September 10, 2019, https://www.forbes.com/sites/bernardmarr/2019/05/20/artificially-intelligent-toilets-yes-they-are-here/#7f3c1585626d.

[3] Neither AI-enabled smart fridges or smart toasters are particularly novel ideas.

[4] Google, Uber and Tesla are some of the large market players working on autonomous vehicles.

remember the times when you used to get stuck in traffic before autonomous vehicles made that a thing of the past. There is something as nostalgic rage after all…

Arriving at work you chat with your colleagues for a few minutes before going to the meeting. Your phone reminds you that Lisa will be there, and it is her birthday. You almost forgot! You wish Lisa a happy birthday and start the meeting on a positive note. After connecting the automatic note taker everyone present may brainstorm at will[5]. Notes will be forwarded to everyone's emails.

Meeting finishes at 1:00 PM so it is already time for lunch. You decide to go to the new restaurant that opened behind your office. It is a new franchise where every restaurant is equipped with a robot chef![6] You can select your food through the app and even suggest new recipes through their website. Payment is wireless and through your account in the app[7]. As you are walking there, you log in and it suggests you a nice honey garlic glazed salmon, which you happily accept.

The app tracks your path through your phone's GPS system and the salmon is waiting for you in your favourite table when you arrive. As you enjoy the delicate dance between the flavours of the salmon, garlic and honey, you think to yourself: "what a brave new world!"[8].

It does seem like a pleasant scenario… AI working along with humanity to make our life easier. But this is hardly the only possible scenario. Our story could as easily start with you waking up to find a barrage of abuse throw at you on Twitter by an AI-enabled bot or fake pictures/videos of you being spread online developed through *deepfake* algorithms. Your vacuum cleaner might try to vacuum your hair in the morning mistaking it for dirt. You could even go out and be shot by an AI-enabled drone that identified you as a terrorist by mistake. We are not even going to real doomsday scenarios, where we are all serving our new and improved robotic overlords.

There is no doubt that, save for any unforeseen circumstance, AI will dramatically impact the way we live. Our conviction and the early signs point to a positive impact, but currently it is not possible to be sure that the final results will follow the trend.

---

[5] Cisco's Voicea offers similar functions to the ones described.

[6] *See*, "NVIDIA's 'kitchen manipulator' is the ultimate robot chef", designboom, accessed September 10, 2019, https://www.designboom.com/technology/nvidia-kitchen-manipulator-robot-chef-cobot-15-01-2019/.

[7] Amazon's brick and mortar stores work on a similar concept.

[8] Our Brave New World is far nicer than Huxley's, though far less brilliant.

Taking this into account it would not be acceptable for the various phases of research, development and deployment of AI to be kept in a legal grey area or to be regulated by legislation that was not built with it in mind. It is possible and productive to have a discussion around what should be regulated and how, and we will certainly partake in it as we go along this Thesis. But question is what and how, not really if. National and supranational entities around the world have identified the necessity for adequate legal tools for AI regulation. Generally, it is acknowledged that AI Regulation should achieve two objectives: *a)* reap the rewards of AI while; *b)* avoiding its pitfalls. As we will see below, this is an enormous challenge.

In particular, the European Union selected achieving a leadership position in the AI sector as one of the priorities for the future of the bloc as a whole. Economic reasons are behind this decision, but we do not think that it is naïve to think that there is something more[9]. Undeniably, AI will have an enormous impact on world's economy and if the EU falls behind, the standards of living that we currently enjoy may be at this risk. However, there is also the question of keeping European values, principles and ethical standards alive in this technological transition. European law-making is imbued with this spirit when acting in regard to new technologies. We will be able to notice this in our analysis of the GDPR and new consumer protection Directives in Part II.

This Thesis is divided in two Parts. Part I is entitled "Efforts for AI and Machine Learning Regulation" and is outward and future focused. In Chapter I we are given the necessary tools to tackle the challenge ahead of us, through some essential preliminary concepts. If Chapter I is the "technical crash course" on what is AI, what are the types of AI and what specific challenges arise from the different types of AI, Chapter II is the "regulating matters in the EU crash course", there we learn how law-making in the EU works and how different choices at this stage may affect the legitimacy and effectiveness of our future regulatory framework.

Chapter III is the European trends chapter. There we will analyse all preparatory documentation issued by the European Institutions and other relevant bodies on AI Regulation. Chapter III has a dual nature, one is descriptive and aims to give the reader a notion of what the European Institutions are planning and what should be expected in the future in terms of AI Regulation (we cannot forget that the preparatory documents are

[9] As we do not think that is naïve to think there is something more on the general digital single market strategy of the EU. *See*, Sophie Perez Fernandes, *"O digitalismo é uma forma de humanismo – o contributo da União Europeia na formatação do humanismo digital como paradigma de vida em sociedade do século XXI",* [forthcoming].

paving the way for it). The second nature of this Chapter is analytical as the positions in the preparatory documentation are frequently criticised and, when possible, alternatives offered.

If Chapter III looks into the body as whole, Chapter V looks into the parts that compose it. What are European Member States doing to regulate and foster AI development? The descriptive nature is maintained, but the analytical one is not unchanged. Focus is less on the mistakes and misjudgements in positions and more on the effects of decisions by Member States on the EU.

Chapter IV leaves home and goes exploring the neighbourhood. What are other countries and economic blocs doing on AI? How are they doing it? Are they being successful? Are our values and ethics aligned?

Part II is entitled "Current EU Legal Framework and Challenges" and it is inward and present focused. In this Part we ask ourselves what legal tools do we already have at the European level that regulate (certain areas of) AI? The EU's body of legislation is enormous and, if we went sector by sector, there would be no end in sight for the legislation that could, in theory, be applicable to AI. Therefore, we chose to centre our work into the GDPR, the Product Liability Directive and the new Consumer Protection Directives. The reason for this choice is that they will be applicable to AI in general (as opposed to sector specific) and they greatly affect the pathway of AI development and implementation in the EU. Theoretically, we also found great interest in the challenges that arose in trying to reconcile AI with the legal dispositions contained within these instruments. Unfortunately, and mainly due to space and time restrictions legislation on matters such Cybersecurity, Intellectual Property, Workers' rights and Medical Devices etc. falls out of our scope.

The three Chapters in this Part follow the abovementioned legal instruments, Chapter I for the GDPR, Chapter II for the new Consumer Protection Directives and Chapter III for the Product Liability Directive.

Finally, we close the Thesis with our Conclusions and some Policy Suggestions for adequate, effective and proper future AI Regulation in the European Union.

PART I – EFFORTS FOR AI AND MACHINE LEARNING REGULATION

Chapter I – Preliminary Concepts

§ 1. Artificial Intelligence

What is artificial intelligence? Indeed, a problematic but pertinent question to start this Thesis. Open any book or essay on the matter and the first chapter will be, almost without exception, about the definition of artificial intelligence. To define artificial intelligence, we must be able to define what is intelligence and that is quite a deep philosophical issue. The "artificial" part of the equation is not as troubling, but even then, we cannot define it just as human-made. If we did, arguably, an intelligence machine that was created by another intelligent machine would not be artificial. A broader and righter definition is that artificial is something that is synthetic, or, in other words, is not naturally occurring. For all the discussion around the issue, a universally accepted definition of artificial intelligence is something that still escapes academics and professionals in the field, even if surely not for lack of trying.

Let us start from the beginning and engage in an extremely brief explanation of what is AI and why are we captivated by it. Human beings had a fascination with the idea of creating artificial human-made intelligent life long before the means to do so were even close to being available. In Greek mythology Hephaestus created the mechanical Talos, a giant bronze automaton built to protect Crete from invaders. Talos was commissioned by Zeus himself, as a gift for his son Minos. In addition, Hephaestus built mechanical women able to cater to his wishes and Pandora, the last with the assistance of Athena and other gods. Meanwhile, Jewish mythology contains the Golem and Buddhist traditions contain stories about the creation of lifelike mechanical doll by an inventor to deceive a respected painter[10/11].

Mary Shelley's "Frankenstein" is probably the most widely known history about artificial human-made life (and its consequences) and is regarded as the first work of science fiction. The genre gained enormous popularity and idea is also present in classics such as H.G. Wells "The Island of Dr Moreau", Karel Capek's "R.U.R", Asimov "I, Robot", Philip K Dick's "We Can Build You" or "Do Androids Dream of Electric

[10] For a plethora of interesting examples *see*, Adrienne Mayor, *Gods and Robots: Myths, Machines, and Ancient Dreams of Technology* (Princeton: Princeton University Press, 2018): p. 7ff.
[11] Arlindo Oliveira*, Inteligência Artificial* (Lisboa: Fundação Francisco Manuel dos Santos, 2019), 37ff.

Sheep[12]" and Douglas Adams "The Hitchhiker's Guide to the Galaxy" series, amongst others. Nowadays, AI (particularly through robots) is ever-present in our books, TV shows, movies and collective imaginary.

In its 1950's "Computing Machinery and Intelligence" paper, Alan Turing suggested the idea of what we now know as AI and designed the Turing Test. According to the Turing Test, a human tester should have a conversation in natural language (absent any constraints) via terminals with both an AI and human. The AI and human are hidden from view. If the original tester is unable to distinguish between AI and human, the AI should be considered as intelligent[13]. Specialists in the field of AI have criticised or proposed amendments to the Turing Test, and it is undoubtedly adapted to general AI and not narrow AI, but it is still an important landmark in the development of the field and used as a metric of AI development.

Modern research on AI arguably began in 1956 when John McCarthy of Dartmouth College, Marvin Minsky of the Harvard University, Nathaniel Rochester of IBM along with Claude Shannon of the Bell Telephone Laboratories organised the Dartmouth Conference. McCarthy was the first to propose a definition of AI, according to which the science of AI is defined as *"the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable"*. Intelligence is defined as "*the computational part of the ability to achieve goals in the world. Varying kinds and degrees of intelligence occur in people, many animals and some machines*", and to the question if AI's objective is to simulate human intelligence McCarthy answers that *"sometimes but not always or even usually. On the one hand, we can learn something about how to make machines solve problems by observing other people or just by observing our own methods. On the other hand, most work in AI involves studying the problems the world presents to intelligence rather than studying people or animals. AI researchers are free to use methods that are not observed in people or that involve much more computing than people can do"*[14]. We should note that McCarthy considered that there was (yet) no solid definition of intelligence that did not depend on relating it to human intelligence. For what is worth, the Dartmouth Conference was responsible for overly optimistic (and

[12] Subsequently adapted into the cult-hit Blade Runer.
[13] *See*, A.M. Turing, *"Computing Machinery and Intelligence"* Mind 49 (1950): 433-460.
[14] "Basic Questions", John McCarthy, accessed February, 5, 2019, http://www-formal.stanford.edu/jmc/whatisai/node1.html.

wrong) predictions regarding the development of AI, such as that it would be widespread in 15 to 25 years.

A problem with McCarthy's definition is that it is excessively broad and, as pointed out by Wolfgang Ertel may classify simple electrical circuits or apparently disorganised robotic vehicles as intelligent. This is mainly due to its weak definition of intelligence. Ertel supports Elaine Rich, Kevin Knight and Shivashankar B. Nair's definition of AI as *"the study of how to make computers do things which, at the moment, people do better*"[15].

While we certainly admit that it is elegant, and will probably age better than other proposals, it is not an adequate definition to build a legal framework upon. Some suggest, and probably rightly so, that, from a scientific standpoint not having a unified definition of AI is actually positive for research and development in the field. Still, we cannot admit that in a legal context, as the legal uncertainty arising from it would be unacceptable[16].

It is possible to define AI coming from two different frameworks, human-centric or rationalistic. As the name implies, human-centric definitions are centred in the human characteristics of AI. Basically, this type of definition states that a machine is intelligent when it is able to carry out tasks that would only be possible for a human being. If we told that an intelligent machine would be one that was able to pass the Turing Test, that would be a human-centric definition. By tying his definition of intelligence to human intelligence McCartey also appears to opt for a human-centric approach, though if not tied to it, his definition of intelligence could probably stand of its own. Another example of human-centric definitions is Ray Kurzweil's "*creating machines that perform functions that require intelligence when performed by people*"[17].

A legal definition of AI must be broad and flexible enough to avoid loopholes arising from unexpected situations. AI legislation must stand the test of time and, thus, will probably be called upon to regulate situations that do not even exist nowadays[18].

---

[15] *See*, Elaine Rich, Kevin Knight and Shivashankar B. Nair's, *Artificial Intelligence*, 3rd ed. (New York: McGraw-Hill, 2009), 3ff.; Wolfgang Ertel, *Introduction to Artificial Intelligence,* 2nd ed. (Cham: Springer: 2011), 1ff.

[16] This is not to say that legal written definitions are always a necessity when regulating a specific field. In his concurring opinion in the US Supreme Court case Jacobellis v. Ohio, Justice Steward concluded, in regards to hard-core pornography that *"I shall not today attempt further to define the kinds of material I understand to be embraced within that shorthand description, and perhaps I could never succeed in intelligibly doing so. But I know it when I see it, and the motion picture involved in this case is not that".* However, we would argue that in AI's case the answer is neither as clear-cut or part of the "general shared knowledge" and people cannot, at the end of the day, be expected to comply with laws that they do not understand. *See*, Judgment of the Supreme Court of the United States of 22 June 1964, *Jacobellis v. Ohio.*

[17] *See,* Ray Kurzweil, *The Age of Intelligent Machines* (Cambridge, MA: MIT Press, 1992), 3ff.

[18] *See*, Iakovina M. Kindyldi, *Smart Companies: Company & Board Members Liability in the Age of AI* (Master's thesis: Tilburg University, 2018), 4ff.

However, there are limitations to this, as AI regulation should be restricted to AI and should not end up regulating other fields of computer science due to an all-encompassing inaccurate and misguided definition. Strangely, human-centric definitions manage to fail at both criteria. On one hand, human beings engage in a lot of activities that are not "intelligent" or, at least, not complex enough to merit special AI regulation. These may be non-brain draining, easily automatable activities such as flipping a light switch or entering into sleep mode when not in use (could be considered as mimicking the human behaviour of shutting down appliance when not in use) to non-logical activities and reactions such tiredness, jealousy etc. It also fails at being sufficiently broad or flexible because they are built based on a wrong preconception: that AI must think like humans or (at least) that human thinking is the best possible thinking. This fails to take into account that, even with narrow AI, for certain tasks, machines have resources that humans just do not. Imagine that you want to produce a dictionary of synonyms based on AI. It is possible to feed every single book written in a language to a machine if sufficient processing power is available. This will not be possible to a human and may be relevant in building our dictionary. Meanwhile, even if the machine is able to compilate a great dictionary, the human will probably still be better at writing a novel, for the foreseeable future. Same for videogames, a machine may be able to play a billion games in a second and will end playing in a manner that is very different from a human[19]. Humans will never be able to do that because our hardware is (in the writer's opinion quite sadly) not designed to do so. Nevertheless, we are still better at designing said videogames.

Rationalistic definitions are goal-oriented definitions. Shane Legg and Marcus Hutter define intelligence as *"an agent's general ability to achieve goals in a wide range of environments"*[20]. Thus, an intelligent machine would be one that is able to achieve goals in a wide range of environments. Nils J. Nilsson considers it to be *"[the] quality that enables an entity to function appropriately and with foresight in its environment"*[21]. One can argue that these definitions fail in that they assume that AI will have an external and static set of goals to

[19] Kamil Rocki a researcher from IBM designed a Gameboy emulator based on deep learning and reinforced learning, able to learn to play games by playing hundreds of games at the same time at 1 billion frames per second. *See*, "A GAMEBOY supercomputer", Kamil Rocki, accessed 15 September 2019, https://towardsdatascience.com/a-gameboy-supercomputer-33a6955a79a4; "This Guy Made a 'Game Boy Supercomputer' That Can Handle 1 Billion Frames Per Second", Daniel Oberhaus, accessed 15 September 2019, https://www.vice.com/en_us/article/qvqamb/this-guy-made-a-game-boy-supercomputer-that-can-handle-1-billion-frames-per-second.

[20] *See*, "A Universal Measure of Intelligence for Artificial Agents", Shane Legg and Marcus Hutter, accessed August 20, 2019, https://www.ijcai.org/Proceedings/05/Papers/post-0042.pdf.

[21] *See*, Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge: Cambridge University Press, 2010), xiiiff..

achieve. For now, and in the narrow AI paradigm, an external source will probably still give the machine an overall high-level objective, as go from point A to point B or find patterns in this dataset. But the AI itself may decide on what steps it takes to achieve the high-level objective, every task is always completed by completing a series of sub-tasks, every goal is realised by realising a series of lower-level goals, and the AI will be able to define at least the lower-level goals and tasks. For a future with general AI (which is out of the scope of this Thesis), we do believe that a definition based a rationalistic definition in this line would be wrong. The difference is that the AI will be able to internally define all its goals, as humans do.

Jacob Turner rejects the notion of both human-centric and rationalistic definitions for use in law-making and proposes, as an alternative, *"Artificial Intelligence is the ability of a non-natural entity to make choices by an evaluative process*"[22]. There are some issues with this definition. First, it only covers ML. The Author does not consider symbolic AI to be a part of ML. Although we do admit that it is not as problematic as other fields of ML, we, in line with Pedro Domingos considered that it could be part of ML. However, even if we did not, AI Regulation should not be reserved for ML, it should also be prepared to regulate issues arising from non-ML AI such as knowledge engineering. As an example, we will study in our chapter about data protection, the issue of the right to explanation and transparency in AI. Even if a black box does not seem as likely in a knowledge engineering-based or in a symbolic AI-enabled agent. The fact is that justifying a decision to a person based upon the existence of a decision tree with 10.000 steps or a rule set with 10.000 rules is as useful as based upon a deep learning model. The person will likely not understand, will not be able to contest and analysing bias or other types of irregularities will still be difficult. The definition may also be interpreted in a manner where it does not cover AI that does not make choices and only supports humans in decision-making.

The Author also considers that "Evaluative Process" should be interpreted as *"one where principles are weighed against each other before a conclusion is reached"* but offers no reason why it should be interpreted in this manner. Is this supposed to be sub-definition? Without an answer to this question, the rest of the argumentation based on Hart-Dworkin debate, as interesting as it is, loses relevance. In addition, since we demonstrated that non-ML AI should also be regulated the definition stands incomplete.

[22] *See,* Jacob Turner, *Robot Rules: Regulating Artificial Intelligence* (Cham: Palgrave Macmillan, 2019), 7ff.

The European Commission proposed a definition of AI as follows: *"Artificial intelligence refers to systems that display intelligent behaviour by analysing their environment and taking actions – with some degree of autonomy – to achieve specific goals"*[23]. This a general rationalistic definition with some shortcomings (as an example, the term specific goals should be avoided due to possible AI developments). We will not perform a deep analysis of this definition because the HLG already updated and perfect it and we are confident that the EC will heed the HLG's suggestions[24]. According to the HLG *"Artificial intelligence systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions.*

As a scientific discipline, AI includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems).[25]*"*

Some immediate red flags are raised by this definition. Although it is more comprehensive than the original it also suffers from excessive complexity and an apparent resolve to be too restrictive. First, the definition is restricted to systems designed by humans. The HLG argues that current systems will always be designed by humans and AI may do no more than support. This will not necessarily be true in the future and, even if it were, this definition would still create a discussion around human design. What level of control does the human need to delegate to a machine before it is no longer human-designed? The definition should, at least, be written in this manner "Artificial intelligence systems are software (and possibly also hardware) systems partially or fully designed by humans". Ideally, the reference to human design would be fully suppressed.

[23] In the Communication Artificial Intelligence for Europe, further analyzed below.
[24] Though after the legislative procedure the final definition may naturally be different from the one currently proposed by both.
[25] *See*, "A Definition of AI: Main Capabilities and Disciplines", High-Level Expert Group on Artificial Intelligence, accessed April 10, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341.

Goals should not need to be given, although we agree that they should be complex. By removing the reference to given goals, we lose nothing in the present and keep our definition adapted to future uses of AI.

AI systems may use symbolic rules or learn a numeric model, or both, but they do not and should be restricted to those two in abstract and by definition.

Lastly, the definition of AI as a scientific discipline is too detailed and should be simplified. With this in mind, we propose the following tweaks to the definition from the HLG: "Artificial intelligence systems are non-natural entities composed by software (and possibly also hardware) that act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve a complex goal. AI systems can use symbolic rules or learn a numeric model, and they may also be able to adapt their behaviour by analysing how the environment is affected by their previous actions.

As a scientific discipline, AI includes several approaches and techniques, such as machine learning, machine reasoning, and robotics"[26].

§ 2. Narrow AI vs General AI

The term Narrow AI, sometimes called Weak AI, refers to a type of AI whose focus is completing specific tasks for which it was developed. If the AI is developed for speech recognition, it is not adequate for written character recognition, if the AI is developed to drive a car it will not pilot an airplane etc. Nevertheless, the framework on the basis of a type of narrow AI can potentially be used for different functions. The same framework could, indeed, learn to drive a car and a bike or learn to play chess and checkers[27]. Every type of AI available today is Narrow AI and albeit limited, it is deeply useful and can help human beings immensely. While it is true that Narrow AI is only able to complete very specific tasks it is quite effective at those and, frequently, better than human beings. Repetitive tasks are especially prone to automation.

[26] *See*, Stuart Armstrong, *Smarter than us: the rise of machine intelligence* (Berkeley: Machine Intelligence Research Institute, 2014): chapter 3.
[27] See below how AlphaZero is able to play chess, shogi and Go.

General AI (or Strong AI) is currently only in the realm of science fiction. The term refers to Artificial Intelligence whose capabilities are akin to those of a human being. Human beings, with all their limitations, are able to perform an almost limitless number of tasks. At the present time, compared to us, AI still has a long way to go. Still, with the quick development of computing power and vast amounts of data General AI is far away, but maybe it is achievable. If the AI surpasses human abilities it is sometimes called an AI Super Intelligence. If we ever reach General AI it is arguable that the temporal gap between it and Super Intelligence will probably be much shorter than the one between Narrow AI and General AI (Narrow AI has already existed for some time, after all). Still, it is not certain that we will ever reach General AI, plenty of scientists and philosophers believe that a machine will never be really able to think like a human. However, for all intents and purposes and keeping with the scope of this Thesis (which is regulating AI) our focus should be on Narrow AI because that is type that exists now and that will exist in the short and medium term and requires an adapted regulatory approach. With this in mind, we should start looking and preparing for the possible development of General AI as to avoid being caught unprepared in what could be a complete shift in our society[28].

§ 3. Software Agents

The concept of software agents has the potential to have a profound impact on issues like regulation of AI-liability and even on data protection (albeit in more a theoretical fashion in this last case).

[28] *See*, Katja Grace et al., *"Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts"* Journal of Artificial Intelligence Research 62 (2018): 729-754; Urs Gasser and Virgilio A.F. Almeida, "A Layered Model for AI Governance" IEEE Internet Computing 21, 6 (2017): 58-62; Seth D. Baum, Ben Goertzel and Ted G. Goertzel, *"How Long Until Human-Level AI? Results from an Expert Assessment"* Technological Forecasting & Social Change 78, 1 (2011): 185-195; Nicolas Miailhe and Cyrus Hodes, *"The Third Age of Artificial Intelligence"* Field Actions Science Reports: The Journal of Field Actions 17 (2017): 6-11; Ben Goertzel, *"Human-level artificial general intelligence and the possibility of a technological singularity A reaction to Ray Kurzweil's The Singularity Is Near, and McDermott's critique of Kurzweil"* Artificial Intelligence 171 (2007): 1161-1173; Jordi Bieger, Kristinn R. Thórisson and Deon Garrett, "Raising AI: Tutoring Matters" in *Artificial General Intelligence: 7th International Conference, AGI 2014, Quebec City, QC, Canada, August 1-4, 2014*, Ben Goertzel, Laurent Orseau and Javier Snaider eds. (New York: Springer, 2014), 1-10; Ben Goertzel and Cassio Pennachin eds., *Artificial General Intelligence* (Berlin: Springer, 2007), 1ff.; Alexandre Veronese, Alessandra Silveira and Amanda Nunes Lopes Espiñeira Lemos, *"Artificial intelligence, Digital Single Market and the proposal of a right to fair and reasonable inferences: a legal issue between ethics and techniques"* UNIO EU Law Journal 5,2 (2019), in Press; Ivan M. Havel, "On the Way to Intelligence Singularity", *in Beyond Artificial Intelligence: Contemplations, Expectations, Applications*, Jozef Kelemen, Jan Romportl, and Eva Zackova (eds.) (Berlin: Springer, 2013), 3-26.

As with AI, there is no standardized definition of software agents. After conceding that we have zero chance of being able to agree on what is a software agent, Hyacinth S. Nwana suggests the definition *"a component of software and/or hardware which is capable of acting exactingly in order to accomplish tasks on behalf of its user"*. The author considers that this is an umbrella term that encompasses such realities as *"knowbots (i.e. knowledge-based robots), softbots (software robot), taskbots (taskbased robots), userbots, robots, personal agents, autonomous agents and personal assistants"*. Nick Jennings and Michael Wooldridge define the phenomena as *"as a self-contained program capable of controlling its own decision making and acting, based on its perception of its environment, in pursuit of one or more objectives"*. Allen Cypher, David Canfield Smith and Jim Spohrer, consider them to be *"a persistent software entity dedicated to a specific purpose"*. Ted Selker opts for *"computer programs that simulate a human relationship by doing something that another person could do for you"*. Tina Balke and Torsten Eymann build upon the definitions of weak and strong agency of Michael Woolbridge and Nicholas R. Jennings[29]. In fact, many more examples of disagreements regarding this definition could be provided. Francisco Pacheco de Andrade who studied the issue deeply also draws attention to the issues around reconciling or at least guaranteeing a certain degree of compatibility between the technical concept of software agents and the concept of agent and representative in legal terms. Though the concept of software agent is extremely developed in his works, when in need to adopt a concise definition Andrade opts for *"software agents are computational entities with a rich knowledge component, having sophisticated properties such as planning ability, reactivity, learning, cooperation, communication and the possibility of argumentation"*. The Author considers that the

[29] Weak agency is characterized by *"autonomy (i.e. the agent's capacity to act without the intervention of its human principal or any other users and thereby having some level of control over its activities and internal states), social ability (i.e. the agent's ability to communicate with other agents and humans through a shared agent communication language), reactivity (i.e. the agent's ability to perceive an environment and respond in a timely fashion to changes that occur within it) and pro-activity (i.e. the agent's ability to demonstrate goal-directed activity by taking initiative). This list of attributes is expanded in Wooldridge's and Jennings' strong notion of agency in which they furthermore mention knowledge, belief, intention, obligation, mobility, veracity, benevolence and rationality as auxiliary characteristics for software agents and thereby attribute software agents all necessary characteristics to, at least virtually, support all stages of the contractual process."*. Detecting an issue with autonomy from a legal standpoint Balke and Eymann add that *"not only one single layer with the software agent that perceives its part of the global environment through sensors and acts upon that environment through effectors needs to be considered, but a second layer with the software agent's principal who has his own perceived environment (e.g. the business context) which he senses and acts upon. The software agent and the principal are in a principal-agent-relationship, meaning that the agent is supposed to act (i.e. negotiate, conclude and carry contracts) on behalf of its human principal in the principal's name"*. See, "The conclusion of contracts by software agents in the eyes of the law", Tina Balke and Torsten Eymann , accessed April 6, 2019, https://www.researchgate.net/publication/221456172_The_conclusion_of_contracts_by_software_agents_in_the_eyes_of_the_law/link/0fcfd50adf322c7aa1000000/download

relevant question to be made will always be if the software agent can represent a natural or legal person or even become a separate legal entity under the law[30].

We also need to distinguish between an agent and a simple object. An adequate distinction seems to be that an object has some degree of control over its own state but not its behaviour. *"When properly addressed (by another object or agent) the object will simply execute the requested task, it has no control over the execution of its methods"*. If you switch on your non-smart television, you will have to use the proper defined methods to access it. But from them on the television has no control over the execution and must respond by showing a picture. By comparison, an ML-software agent will have control over how it executes a certain task. You may tell it to distinguish between cats and dogs, but you will not have full control over how it is done, in fact, it will learn thousands or millions of rules that that it will use in the future to accomplish the task at hand and the human-programmer may not be fully aware or fully understand the rules behind the result, even after looking at the source code[31].

What is then the relationship between AI, ML and software agents? Well, ML algorithms will be software agents. In fact, the whole objective of using ML techniques it to build better and more advanced software agents that are able to help in accomplishing tasks or accomplish those said tasks by themselves. Software does not have to be based on ML and can be instead based on AI techniques such as knowledge engineering. ML-derived software agents are, arguably, the most advanced type of software agents and, thereby, may create challenges that did not arise with less complex and older types of software agents. For the purposes of this Thesis we will consider that autonomous robots are also (powered by) software agents, even though they interact with the world through sensors and other input devices and, in principle, will respond to stimulus from the external world. On what concerns this specific issue we would point out that every autonomous Robot needs to have in its basis one or more software agents working in a coordinated

[30] The Author's rationale is applied to electronic commerce, though we think that it is applicable in general. Furthermore, Francisco de Andrade considers software agents exclusively as software (excluding robots). Even though the Author himself points out that some considerations will also have applicability in robotics we must point out that it is our opinion that, any autonomous robot will have one or more software agents working in a coordinate manner incorporate into its design. Therefore, we prefer to no restrict the concept, in line with Hyacinth S. Nwana's. *See*, Francisco Pacheco de Andrade, "Da Contratação Electrónica – Em Particular da Contratação Electrónica Inter-Sistémica Inteligente" (PhD Diss., University of Minho, 2008), p. 159ff.; ;."Defects of the will in software agents contracting", Francisco de Andrade, et al, accessed May 10, 2019, http://repositorium.sdum.uminho.pt/handle/1822/19096,

[31] *See*, B.W. Schermer, *Software agents, surveillance, and the right to privacy: a legislative framework for agent-enabled surveillance* (Leiden: Leiden University Press, 2007): 17-35.

manner incorporate into its design. The software agent can be developed through the use of ML (or not). Without said software agent(s) as its brain the robot would be no more than a collection of nuts and bolts[32].

It is important to understand (at least broadly) the concept of software agents since it could have some relevant on liability (even if its core relevance is not on Product Liability and more on other types of contractual and non-contractual liability).

§ 4. Machine Learning and Knowledge Engineering

Knowledge Engineering is a field of artificial intelligence where knowledge is introduced directly into the computer by human expert programmers. If we wanted to create an artificial human through the magic of knowledge engineering we would have to program all the functions of the body, from some that appear to be more or less straightforward such as making a beating heart, to others that are currently unthinkable such as programming our senses or a fully functional brain.

---

[32] *See*, Francisco Pacheco de Andrade, "Da Contratação Electrónica…, p. 157-285.; Francisco Pacheco de Andrade et al., *"Contracting agents: legal personality and representation"* Artificial Intelligence and Law 15, 4 (2007): 357-373; Francisco Pacheco de Andrade, Davide Carneiro and Paulo Novais, *"A inteligência artificial na resolução de conflitos em linha"* Scientia Ivridica: Revista de Direito Comparado Português e Brasileiro 59,321 (2010): 1-28; Francisco Pacheco de Andrade, José Neves and Paulo Novais, "Software Agents and Contracts", *in Encyclopedia of Networked and Virtual Organizations,* Goran Putnik and Manual Cunha eds., (Pennsylvania: Idea Group Reference, 2008), 1-7; Francisco Pacheco de Andrade et al., *"Software Agents and Virtual Organizations: Consent and Trust"* International Journal of Services and Operations Management 6,3 (2010): 352-361; Francisco de Andrade et al., "Agents, Trust and Contracts", *in* Irene Maria Portela and Maria Manuela Cruz-Cunha (ed.), *Information Communication Technology Law, Protection and Access Rights: Global Approaches and Issues* (Hershey : IGI Global, 2010): p. 188-199;."Defects of the will…", Francisco de Andrade, et al, accessed May 10, 2019, http://repositorium.sdum.uminho.pt/handle/1822/19096, "An agent-based approach to consumer´s law dispute resolution" David Rua Carneiro et al., accessed May 10, 2019, https://repositorium.sdum.uminho.pt/handle/1822/19079; Jeffrey M. Bradshaw "An Introduction to Software Agents", *in* Jeffrey M. Bradshaw (ed.), *Software Agents* (Cambridge: MIT Press, 1997), 3-46; Hyacinth S. Nwana, *"Software Agents: An Overview"* The Knowledge Engineering Review, 11,3 (1996) 205-244; Hyacinth S. Nwana and Divine T. Ndumu, "A Perspective on Software Agents Research" *The Knowledge Engineering Review* 14,2 (1999): 125-142; Nick Jennings and Michael Wooldridge, *"Software Agents"* IEE Review 42,1 (1996): 17-20; Pedro Miguel Freitas, Francisco Andrade, and Paulo Novais, "Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible", in. *AI Approaches to the Complexity of Legal Systems: AICOL 2013 International Workshops, AICOL-IV@IVR Belo Horizonte, Brazil, July 21–27, 2013 and AICOL-V@SINTELNET-JURIX, Bologna, Italy, December 11, 2013, Revised Selected Papers,* Pompeu Casanovas et al (Heidelberg: Springer, 2014): p. 145-156; Francisco de Andrade et al., "Software Agents As Legal Persons", in Luis M. Camarinha-Matos (ed.), *Virtual Enterprises and Collaborative Networks: IFIP 18th World Computer Congress TC5 / WG5.5 - 5th Working Conference on Virtual Enterprises 22–27 August 2004 Toulouse, France* (Heidelberg: Springer, 2004), 123-132; David Canfield Smith, *"Programming Agents without a Programming Language"* Communications of the ACM 37,7 (1994): 55-67; "The Validity and Limitations of Electronic Agents in Contract Formation: A brief discussion under the Electronic Transactions Act in Australia", Adrian McCullagh, accessed April 6, 2019, https://law.uq.edu.au/files/18238/A-McCullagh_The-Validity-and-Limitations-of-Software-Agents-in-Contract-Formation.pdf.

Originally, knowledge engineering was the preferred path by computer scientists for the development of AI. In the early years, data was not as widely available as it is today, and neither was the computing power to process and learn from such data. Therefore, it is natural that coding knowledge directly into a computer would seem like the most viable path to achieving intelligence machines. However, knowledge engineering was unable to leap to the next level due to some innate limitations. The most important is that, eventually, the program created by introducing manually information into a computer gets so complex that when bugs appear it is impossible to fix them. There are too many parts to the equation and human programmers are, at a certain stage, unable to properly deal with them. Currently, extracting knowledge from data is much cheaper than paying people to both acquire and introduce it into a computer compatible format. Purely knowledge engineering-based software and hardware would also need to be updated constantly, as it lacks the ability to learn things by itself. Imagine you build a knowledge engineering-based medical equipment for detecting cancer in patients and you introduced into it every state-of-the-art approach and technique. Your device will only be up to date until the next issue of Nature or The Lancet as new techniques appear and then you will have to understand them and introduce them manually in the device since it is not able to learn by itself. That is another innate limitation of knowledge engineering, it can know as much as the best humans, but it can never know more.

Due to these limitations, ML took the front seat in AI development in the last few years and it is unlikely that knowledge engineering ever gets its crown back. Presently, most literature regarding AI is actually about ML. The need to study new and specific AI regulation was also brought by the rapid advancements due to machine learning technology and it is clearly the main regulatory priority. This Thesis is no exception and our scope is, in fact, regulation of AI with a specific focus on machine learning, from where the most important advancements and challenges in the near future are likely to come. Therefore, when we talk about "AI" we are talking about machine learning unless the contrary is explicitly stated. That being said, proper AI regulation should also cover other types of non-ML AI, including knowledge engineering and (as we already had the opportunity to do so, above) when possible we will try to ensure that our suggestions are universal and can be applied to both paradigms.

What is ML then? ML is the currently dominating paradigm on AI development and has been enjoying growing and continuous success on the back of wider availability of data and more powerful hardware that permits processing of the abovementioned data.

Machine Learning AI key characteristic is that it is able to learn by itself, that is, it possesses self-learning capabilities. It is able to learn more and even improve itself based on new information (data) provided to it or acquired through real-world interaction. Thus, unlike with knowledge engineering, to develop AI-based on ML techniques one does not have to program all the necessary knowledge into the machine, allowing ML to overcome major pitfalls that hindered knowledge engineering. The specific method through which ML AI learns depends on which ML technique is used to build a certain algorithm. Although there is a common element to them, the challenges presented by different ML schools are diverse enough to justify having brief knowledge of each of them. In this manner, we will be able to adapt our legal solutions to the specific challenges if needed and pay close attention to the most problematic ones. Thereby, in the following chapter we will briefly look into the major schools of ML.

§ 5. Schools of Machine Learning[33]

§ 5.1. Symbolic AI

Symbolists are the school of machine learning whose beliefs more closely resemble the ones held by knowledge engineers. In fact, not all symbolic AI can be considered to be a part of machine learning, some sits closer to knowledge engineering. In fact, it is frequent to consider symbolic AI separately or even in opposition to machine learning. However, for techniques such as inverse deduction it seems fair to categorize them under machine learning and consider symbolic AI as one of the schools of machine learning.

Due to its early dominance in the field, symbolic AI can also be known as GOFAI (Good Old-Fashioned AI). Symbolic AI has, indeed, been a staple for quite some time. Some early successes based on it, like Arthur Samuel's checkers player (who was able to beat its creator, though it was surpassed in the 70s), Allen Newell, Herbert A. Simon and Cliff Shaw's Logic Theorist able to prove theorems contained within Principia Mathematica (and even to find more elegant proof for one of them) and the General Problem Solver.

[33] In Section 5 we shall fall closely the teachings of Pedro Domingos in *The Master Algorithm*. See, Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World* (New York: Basic Books, 2015). Some examples, where not specifically stated, are from the abovementioned book.

For symbolists, intelligence and symbol manipulation are one and the same. In the same manner that a mathematician solves equations using symbols or a writer writes words, sentences and, ultimately, books using different sets of symbols, an algorithm can also through symbol manipulation achieve intelligence. Currently, the Symbolists' inverse deduction (induction[34]) – including decision tree induction – is used in applications like Microsoft's Kinect. In fact, in a 2002 competition, decision trees beat a panel of experts in guessing decisions from the Supreme Court of the United States (75% to 60%). However, the symbolistic approach can be very resource-intensive and have problems with scalability, is vulnerable to noise and has trouble working with imprecise concepts.

§ 5.2. Evolutionaries

The mere fact that this Thesis exists is proof that evolution served us good. The Author could take the time to write because his more immediate needs like food and shelter are covered. A plethora of investigative resources are also available at the tip of our fingers or after a trip to the library (said trip is not always as short as it would be ideal since there is large – an arguably undesirable – difference in availability between different locations).

Evolutionaries want to apply the theory of evolution to machine learning. Their method in short: something like a random assembly of algorithms is given a certain task[35]. Most likely, at first, they will not be effective at said task. Nevertheless, some will probably be better (or less bad) than the others at it. Those will be selected through a fitness function and "keep living". After the initial culling random "mutations" and crossover[36] will be inserted in the algorithms. If the mutations are benefic for the desired task the algorithm will achieve a better overall score and "survive" for the next round. In the end, we will (hopefully) get an algorithm that will be incredibly good at achieving the task for which it is intended. Competition between algorithms (the dynamics of the predator-prey relationship) can also be introduced to simulate conditions found in natural evolution. Programmers registered positive results with this strategy, making the algorithms evolve much faster when faced with "artificial" competition.

---

[34] Inverse deduction may be used to teach machines to automatically fill gaps in knowledge by use of inductive methods.

[35] The task will be something like producing a model that allows them to distinguish between different breeds of cats.

[36] Crossover is based on the way the genetic makeup of both parents combines to create the genetic makeup of their children in sexual reproduction.

Our evolution was quite slow, the oldest *homo sapiens* fossils ever discovered are from an ancestor that lived more than 300.000 years ago[37]. However, the first cultivated fields date from 13.000 BC and computers are children of the XX century[38]. That is to say, we took some time to get where we are today, and any algorithm based on evolution would not be viable if it worked at the speed of "classical" evolution. Luckily, you can play Civilization or Age of Empires faster than real human evolution and, in the same manner, a sufficiently powerful computer can simulate the evolution of algorithms at a much more acceptable speed. In 2005 the US Patent and Trademark Office awarded a patent to a genetic algorithm.

There are some limitations to genetic algorithms: they are prone to converge to local optima instead of the global optimum, are unable to acquire existent information and may suffer from transparency problems (we will get to that in Part II, our chapter about the GDPR). Still, the combination of genetic algorithms and connectionist algorithms like neural networks and its subset deep learning shows promise.

§ 5.3. Connectionists

The human brain is an impressive and highly refined structure and, at least until now, no machine has come even to close to achieve its level of intelligence. That is to say, a calculator may be better at math than a human, but the machines' versatility pales in comparison to what is possible for the human brain[39].

Therefore, connectionists, through neural networks aim to emulate the human brain's functioning. Chips cannot compete with the number of connections that our neurons are able to achieve. However, they do fire significantly faster than our neurons, (even though our brains using about the same power of a small lightbulb are significantly more effective).

Training a supervised neural network may look something like this: input data is fed to the input layers, while the solution is "given" to the neural network, so it knows what the results in the output layers have to be. The hidden layers will create the necessary

---

[37] "Homo sapiens 100,000 years older than thought", Clive Cookson, accessed April 12, 2019, , https://www.ft.com/content/00a266c2-4ad3-11e7-919a-1e14ce4af89b

[38] *See*, "The Slow Birth of Agriculture", Heather Pringle, accessed April 12, 2019, https://science.sciencemag.org/content/282/5393/1446

[39] Let us not forget that the complete saying is not Jack of all trades, master of none. It is Jack of all trades, master of none, though oftentimes better than master of one.

rules to achieve the desired result and then the model produced can be applied in real-world settings. Of course, neural networks, after being deployed, have the potential to adapt continually to their use.

A particularly successful subset of neural networks is deep learning. In deep learning, networks are built with a high number of hidden layers, where the *"first hidden layer becomes the input/layer of the second and so on"*. Through this technique each *"each hidden layer learns a more sophisticated representation of the input, building on the previous one"*[40].

Deep learning is currently the most successful type of machine learning but also one of the most prone to creating black boxes, as what happens in the hidden layers can be imperceptible to the user and even to the programmer. Notable examples of deep learning use are self-driving cars, tools for diagnosing in medicine, machine translation and image and voice recognition amongst others[41/42].

§ 5.4. Analogizers

Analogical reasoning is the basis of some of the greatest scientific discoveries in the world such as Darwin's natural selection or Bohr's model of the atom. On that note, analogical reasoning is also one of the key parts of legal interpretation and key for filling legal loopholes. Space was for a number of years largely regulated by maritime law, using analogical interpretation. Analogy is, as we can see, quite a useful and flexible tool.

Analogizers enjoy a respectable degree of success in AI with algorithms like the nearest neighbour, support vector machines and algorithms for case-based reasoning. Nearest neighbour is one the quickest and more effective algorithms ever invented, and it works in quite a simple manner. Imagine you want to build an entire Social Network dedicated to sharing pictures of cats (truly the noblest of efforts!) and your users can tag their cats and their breeds. Suddenly you find yourself with a huge dataset containing millions of pictures of cats tagged with their breeds. You then decide to make the tagging of breeds not mandatory but optional. Some of your users are more prone to laziness and thus decide that they should only upload the picture of their cat without the breed. How

[40] *See*, Pedro Domingos in *The Master Algorithm*…117ff.
[41] *See*, Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence* (New York: Alfred A. Knopf, 2017), 71ff.
[42] *See*, "Top 15 Deep Learning applications that will rule the world in 2018 and beyond", Vartul Mittal, accessed June 10, 2019, https://medium.com/breathe-publication/top-15-deep-learning-applications-that-will-rule-the-world-in-2018-and-beyond-7c6130c43b01.

can you ensure that pictures in your social network are still meticulously organized even without the users' manual labelling? Simple, you run your nearest neighbour algorithm and it finds the picture that more closely resembles that particular cat and tags it as so. In the future you can use the model[43] created to tag similar situations. If it looks like a Maine Coon it is probably a Maine Coon. Of course, the success of this method depends on you having a previous dataset to compare it to, the bigger the better. Unfortunately, nearest neighbour has issues with scalability and trouble with separating relevant from irrelevant information.

Support vector machines surpassed nearest neighbour as the most used analogy-based algorithm in the 1990s. Support vector machines are able to resist particularly well to the issues that plague other algorithms like overfitting and scalability. Early successes of this type of algorithm were registered in text classification and the recognition of handwriting. Connectionists have a long-standing feud with analogizers, tough currently the former have the upper hand due to the effectiveness of deep learning algorithms.

These types of algorithms can and have been applied to highly complex matters like customer service, music composition and the law. In fact, an algorithm based on analogical reasoning was able to "guess" the result of 90% of the trade secret cases brought before it.

§ 5.5. Bayesians

Thomas Bayes was an English mathematician and theologian who is responsible for inventing the Bayes' Theorem (even though Laplace was the one who rediscovered it). The Bayes theorem is particularly useful when dealing with situations that lack certainty

[43] The difference between an algorithm and a model is purely technical and not necessarily consistent across works in this field. Therefore, and due to the fact that it does not affect the scope of our Thesis, which is about regulating Artificial Intelligence, we shall base our concepts on the ones put forward by the Norwegian Data Protection Authority (Datatilsynet) in its Artificial Intelligence and Privacy report, along with the definitions contained in Amazon's Machine Learning Developer Guide. We will use the first definition because it comes from a data protection supervisory authority within the European Union and thus, absent a definition contained in hard law, should be seen as the closer we have to a "soft law" instrument on the matter. The reason for using completing it with the second is because Datatilsynet's definition by itself would not be complete enough, Amazon is a leader in ML development, and it is clear and simple enough while being technically accurate. Still, to avoid constant sentences like "the ML model created by the ML algorithm" we will frequently use the term "algorithm" to encompass the entire procedure from learning to result, as in to model that when inputted new situations can make a prediction. *See*, "Artificial Intelligence and Privacy", Datatilsynet, accessed August 15, 2019; https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf "Amazon Machine Learning: Developer Guide", Amazon, accessed August 15, 2019, https://docs.aws.amazon.com/machine-learning/latest/dg/machinelearning-dg.pdf#types-of-ml-models.

this theorem allows for *"on the basis of evidence -- the best (human) estimate that a particular event will take place. You start with an initial estimate of the probability that the event will occur and an estimate of the reliability of the evidence. The method then tells you how to combine those two figures -- in a precise, mathematical way -- to give a new estimate of the event's probability in the light of the evidence. In some highly constrained situations, both initial estimates may be entirely accurate, and in such cases Bayes' method will give you the correct answer. In a more typical real-life situation, you don't have exact figures, but as long as the initial estimates are reasonably good, then the method will give you a better estimate of the probability that the event of interest will occur"*[44].

The mathematical representation of the theorem is the following:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

P(A | B) is the likelihood of the event A occurring when B does.

P(B | A) is the likelihood of B occurring when A does.

P(A) and P(B) are the odds of A or B happening in an independent manner.

We shall give a (very simple) example. Snoopy, the dog, is looking at the window and notices that dark grey clouds are forming on the horizon. Dark grey clouds sometimes bring with them thunderstorms and Snoopy is quite afraid of them. How can Snoopy calculate the probability that a thunderstorm will indeed happen? Snoopy knows that, in April, the probability of a thunderstorm happening is 10%, and the probability of dark grey clouds appearing in the sky on any given day sits at 40%. The probably of the sky being filled with dark grey clouds when there is a thunderstorm is 50%.

$$0{,}125 = \frac{0{,}5 \; x \; 0{,}1}{0{,}4}$$

The probability of thunderstorm is 12,5%. Snoopy will, most likely, be safe.

[44] *See*, "The legacy of the Reverend Bayes", Mathematical Association of America, access March 18, 2019, https://www.maa.org/external_archive/devlin/devlin_2_00.html.

Another example could be given with the test for a certain disease. If a certain person takes the T-Virus test in Portugal and gets a positive result, what is the probability of that person having the T-Virus? Let us imagine that the T-Virus is a very rare disease and that only affects 0,7% of the population[45]. The test will fail 5% of the time and get it right the other 95%. What is the probability of the person having said disease? 95%? 90%? It may seem like it, but not really, it is quite lower.

$$0{,}133 = \frac{0{,}95\ x\ 0{,}007}{0{,}05}$$

The probability is no more than 13,3%. More likely than not, said person is healthy. Of course, it is also possible to represent more than two variables with Bayes theorem, but we shall leave that to technical (not legal) works.

For Bayesians this theorem should be in the basis of ML and indeed, it has found some success in spam filters, search engines (Google's PageRank derives from this type of algorithm), advertising (Google's Adsense uses it), speech recognition (used by Apple's Siri) and Gaming (Xbox live ranks and pairs players on the basis of a Bayesian algorithm). Specific algorithms of the Bayesian family include Naïve Bayes, Markov Chains, Hidden Markov Models and Bayesian Networks. However, the Bayesian approach is not as flexible as, for example, Symbolic AI. Computational requirements and prior probability distribution are other known issues[46].

§ 6. Supervised, Unsupervised, Semi-Supervised and Reinforced Learning

§ 6.1. Supervised Learning

In supervised learning, a training set of data is given to the algorithm. The training set contains input-output pairs of whatever the programmer is teaching the algorithm. For example, if we want to teach our algorithm to distinguish between Sphynx and Persian Cats, we provide the algorithm with a training set that contains as many photographs of those two breeds as we can muster, along with the correct classification. The algorithm

[45] With Leon on the case, Umbrella Corp. will not prevail!
[46] See, Pedro Domingos; "Thomas Bayes", Encyclopedia Britannica, accessed March 17, 2019, https://www.britannica.com/biography/Thomas-Bayes

will then (hopefully) recognize that certain attributes are present for each of the breeds, *i.e.* Persian cats have long fur and flat snouts while Sphynxes have little to no fur and longer snouts. The algorithm (specifically the model created through the previous data) will be then given a second set of data, in our case: pictures of Sphynx and Persian cats and will have to use what it learned to distinguish between the two breeds itself.

Learning to distinguish between breeds of cats is indeed a noble and humanity changing effort. We are quite sure that millions of Youtube users would be delighted by the ability to sort their cat videos playlist by breed, but for the sake of argument let us present another (hypothetical) example of supervised learning.

Credit card fraud is a sizable problem for the Banking Sector, with projected losses of 12 billion[47] dollars[48]. Our algorithm should be able to detect credit card fraud, more precisely, credit card cloning, therefore we give it a dataset containing both legit credit card transactions and fraudulent ones. The algorithm will (hopefully) understand the underlying patterns that are common for the use of clone cards.

Quality of data is obviously key in Supervised Learning, otherwise our algorithm will learn incorrect patterns and potentially overfit. Overfitting happens when a model absorbs irrelevant our outright harmful patterns from our training data, turning it ineffective when confronted with any other set of data. If every transaction from India in our training set and test set were classified as fraudulent then, when applied to real life, our algorithm would probably produce a model that would lock the card of our Indian customers (and customers travelling to India)[49].

Examples of algorithms where supervised learning can be used are neural networks, decision trees and support vector machines.

---

47 For context, we are using the short scale of Billion in which 1 Billion is the equivalent to 1 thousand million (i.e. 1,000,000,000).

48 *See*, "U.S. Card Fraud Losses Could Exceed $12B By 2020", Roger Aitken, access January, 23, 2019, https://www.forbes.com/sites/rogeraitken/2016/10/26/us-card-fraud-losses-could-exceed-12bn-by-2020/#276281c6d243.

49 *See*, Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach"*, (3.ª Ed, Essex: Pearson Education Limited, 2016), 695ff.; Jerry Kaplan, *Artificial Intelligence: What Everyone Needs to Know*, (Oxford: Oxford University Press, 2016), p. 30-32; Richard E. Neapolitan and Xia Jiang *Artificial Intelligence: With an Introduction to Machine Learning,* (CRC Press: Florida, 2018): p. 89ff.; Lior Rokach and Oded Maimon, "Supervised Learning", in *Data Mining and Knowledge Discovery Handbook*, Lior Rokach and Oded Maimon eds., (2.º Ed, New York: Springer, 2010): 133ff; Mireille Hildebrand, "Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning", Theoretical Inquiries in Law 20,1 (2019): 83-122.

§ 6.2. Unsupervised Learning

It is not possible and sometimes not even desirable to have previous examples of labelled data at the disposition of the algorithm. In our current world with the proliferation of large quantities of unlabeled data (big data), the ability to learn from this data without any "ground truth to compared it to" gained the utmost importance. Unsupervised learning can be especially important in areas like pattern recognition, market basket analysis, web mining, social network analysis, information retrieval, recommender systems, market research, intrusion detection, and fraud detection.

Unsupervised learning can be distinguished from supervised learning by the fact that in the first one, there is no previous labelling of the data available to train the algorithm. Arguably the most well-known and used method of unsupervised learning is clustering. In clustering what the algorithm does is separate the data is groups or clusters according to the characteristics it finds. Clusters can then be analyzed, and relevant information can derive from it. Patterns discovered by unsupervised learning can also be used to create datasets used in supervised learning.

We shall go back to our previous Sphynx and Persian breeds. While it is true that without our previous indication that this is Sphynx, and that is Persian our algorithm would probably not be able to define the further cats it encounters as such. However, it would still be able to find differences between Persians and Sphynxes (the lack or abundance of fur, the snot etc.) and probably group the photos of the two. This would particularly useful if we had never seen those cats before.

Regarding our credit card fraud example. We do know some patterns of credit card cloning, but we certainly do not know all patterns. Data scientists actually found that if a credit card was used to buy a gallon of gas, it was an indication that it was stolen (thieves usually did this to check if the credit card was in working condition). Therefore, by giving large quantities of raw data to unsupervised learning algorithms we can use the fact that computers are highly effective at finding patterns, to discover new ones. In our case, we can discover new behavioural patterns from people who clone or use cloned credit cards and with those heighten the security of our banking sector and prevent further losses. Patterns discovered using unsupervised learning algorithms can even function as the basis

for datasets used to teach supervised learning algorithms. Of course, unsupervised learning algorithms can also overfit and start "dreaming" about patterns that are not really there[50].

§ 6.3. Semi-Supervised Learning

Somewhere in between supervised and unsupervised learning, sits semi-supervised learning. As the name implies, semi-supervised learning uses both labelled data (supervised learning) and unlabeled data (unsupervised learning) to teach the algorithm. The algorithm is provided with some input-output pairs but not for all (or most) of the data.

Going back to our Persians and Sphinxes examples, the algorithm can use the labelled data to identify the specific breeds of cats in the data and what they might be. After that, the unlabeled data will help define the boundaries of what constitutes each breed and it might even identify new breeds with are not similar to the previous examples given to it by the labelled dataset, like Siamese cats.

Semi-supervised learning is usually seen as a type of supervised learning with additional information on the distribution of the examples, but certain types of semi-supervised learning can also be seen as unsupervised learning with constraints.

Semi-supervised learning can, in certain cases, achieve better results than supervised and unsupervised learning and offers cost-advantages against supervised learning since manually labelling data is time-consuming and expensive. Furthermore, bias from the human responsible for the labelling can be injected in the data, which the large unlabeled dataset can minimize[51].

---

[50] *See*, David Danks, "Learning", in Keith Frankish and William M. Ramsey eds, *The Cambridge Handbook of Artificial Intelligence* (Cambridge: Cambridge University Press), 151-167; Amparo Albalate and Wolfgang Minker, *Semi-Supervised and Unsupervised Machine Learning: Novel Strategies* (ISTE and Wiley: Chippenham and Eastbourne, 2011): pp. 15ff.; "Unsupervised Learning", Zoubin Ghahramani, access February 15, 2019, http://mlg.eng.cam.ac.uk/pub/pdf/Gha03a.pdf; "Unsupervised Learning", Peter Dayan, access February 16, 2019, http://www.gatsby.ucl.ac.uk/~Dayan/papers/dun99b.pdf; M. Emre Celebi and Kemal Aydin (eds.), *Unsupervised Learning Algorithms*, (London: Springer, 2016), 3ff.; Richard E. Neapolitan and Xia Jiang *Artificial Intelligence*…,p. 331ff.; *Artificial Intelligence: What Everyone* …, p. 30-32; Pedro Domingos, *The Master Algorithm*…, 68-70.

[51] *See*, "What is Semi-Supervised Learning?", Nikki Castle, access February 20, 2019, https://www.datascience.com/blog/what-is-semi-supervised-learning; Olivier Chapelle, Bernhard Schölkopf, and Alexander Zien (eds.), *Semi-Supervised Learning* (Massachusetts: MIT Press, 2006): p. 1ff.; Stuart J. Russell and Peter Norvig, *Artificial Intelligence*…, 695ff.

§ 6.4. Reinforcement Learning

Mr S. has recently moved to a new house and he decided that, for health reasons, he would start walking every morning to his job. There are five possible paths that Mr S. can take and to choose one Mr S. gave priority to three aspects: *i)* he wants to arrive safely; *ii)* he does not want to be late and; *iii)* he wishes for a picturesque view. Mr S. will probably try each path and start taking it more frequently if it lives up to his expectations or avoid it if something unpleasant happens. If Mr S is mugged or arrives late it is a negative point for the path he selected that day and he will be less likely to select it in the future. If Mr S. likes a historic building that is in the path or if he arrives early, he will be more likely to take that path again. Eventually, Mr S. will probably start taking only one path. In a nutshell Mr S. will be doing something very similar to reinforcement learning.

In reinforcement learning indications are not given about what to do, but instead the algorithm must try different combinations, and either be rewarded or punished by them, thus choosing to repeat or avoid those actions in the future. The reward or punishment can be no more than a numeral value associated with a determined outcome. Actions that produce positive feedback are more likely to be repeated than actions that produce negative feedback.

Due to this trial and error nature reinforcement learning faces a problem that does not affect the other learning paradigms: the trade-off between exploration and exploitation. When the algorithm identifies a course of action that results in rewards should it keep repeating or explore new courses of action that might bring it bigger rewards but that do not offer any guarantees of success. If Mr S. is satisfied with the 2nd path should he still try to take the 3rd, 4th and 5th paths in the hope that they are even more pleasurable or just keep with the 2nd path that is guaranteed to bring positive results. The algorithm may be, thereby, stuck with an answer that is good, but not the best possible (or even one of the best).

There is still no solution for the problem, but it is being laboriously studied and its existence does not mean that reinforcement learning is not useful, in fact, far from it. DeepMind's[52] AlphaZero – the successor of Alpha Go, the algorithm that beat the legendary Go player Lee Sedol in front of an audience of 200 million people and won a 9

[52] Google bought DeepMind in 2014 for $500 Million. *See*, "Google Acquires Artificial Intelligence Startup DeepMind For More Than $500M", Catherine Shu, access March 20, 2019, https://techcrunch.com/2014/01/26/google-deepmind/.

dan professional ranking in the process – learned how to play 3 different games (Go, Chess and Shogi) at a above human level[53] in just three days, using a combination of deep learning, reinforced learning and symbolistic AI. The uses for reinforcement learning, obviously, are much broader than playing games with applications that can go from robotics to medicine and medical treatments and personalized recommendations[54].

§ 7. Black Box Algorithms

Systematically, the issue of black box algorithms could be placed almost in every chapter of this Thesis. Black box algorithms are an enormous challenge for personal data protection (probably the area where they have got the most attention, possibly due to the fact that the GDPR is a "fashionable" legal instrument)[55]. We will talk plenty about black boxes in our chapter on the GDPR, including about their relationship with the right to information and explanation. However, there is more to life than just personal data and black boxes also arise in situations where non-personal data is processed.

They are a challenge for Product Liability, where proving either a "defect" or using one of the legally enshrined defences may become much more difficult under black box conditions.

We could also explain black box algorithms under our chapter on European legal initiatives on AI, or even Member State-led initiatives. There is basically no national

---

[53] AlphaZero actually beat AlphaGoZero which had beaten the original AlphaGo. *See*, David Silver et. al., *"A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play"* Science, 362,6419 (December 2018): 1140-1144; "Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm", David Silver et al., access March 15, 2019, https://arxiv.org/pdf/1712.01815.pdf; David Silver et al., *"Mastering the game of Go with deep neural networks and tree search"* Nature 529 (January 2016): 484-507; " The story of AlphaGo so far", DeepMind, access March 15, 2019, https://deepmind.com/research/alphago/; "Move over AlphaGo: AlphaZero taught itself to play three different games", Jennifer Ouellette, access March 15, 2019, https://arstechnica.com/science/2018/12/move-over-alphago-alphazero-taught-itself-to-play-three-different-games/.

[54] *See,* Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction"* (Massachusetts, MIT Press, 2018), 1ff.; Jens Kober, J. Andrew Bagnell and Jan Peters, *"Reinforcement learning in robotics: A survey"* The International Journal of Robotics Research 32,11 (2013): 1238-1274; "Artificial Intelligence: What Is Reinforcement Learning - A Simple Explanation & Practical Examples", Bernard Marr, access January 10, 2019, https://www.forbes.com/sites/bernardmarr/2018/09/28/artificial-intelligence-what-is-reinforcement-learning-a-simple-explanation-practical-examples/#59db807f139c.

[55] Replicating the public interest around the GDPR in any future AI regulation will be very important to ensuring public trust in said regulation. Ensuring that the referred public interest appears while the legislative procedure is ongoing (the ideal time) can be achieve by focusing on citizen participation and having a transparent legislative procedure (see our chapter on the legislative procedure in the EU).

strategy from Member States, Communication from the Commission or deliverable from the HLG that does not touch on the problem of non-explainable algorithms.

However, for a matter that will be transversal across this work, setting up a preliminary explanation seems like a good policy. The challenges brought by black box algorithms and how they can be solved will be then developed in the adequate chapters.

Simply put, there is a black box when humans are unable to predict and/or understand the AI's decision-making process and its outputs. According to the Villani Report (see our section on the French National AI Strategy for more information) they occur when *"it is possible to observe incoming data (input) and outgoing data (output) in algorithmic systems, but their internal operations are not very well understood"*[56].

Some black boxes may be "stronger" than others, as depending on their origin they may be completely opaque to humans or be prone to some degree of reverse engineering and understanding, albeit with considerable effort. Of course, a top-class ML expert will be able to probe much more from a black box than the end-user or even the end-deployer of the AI technology. Since ML experts are highly sought and, thereby, probably unavailable to analyse each "peculiar" decision by an AI, some weaker black boxes, in theory, may end up being impenetrable in practice.

Some families of machine learning are highly prone to creating black boxes that will be very difficult to crack. Evolutionary or algorithms based on analogy are an example. Alarmingly, the current star of ML, Deep Learning, is extremely disposed to to creating black boxes as spectacular as the results it achieves[57].

[56] *See*, "For a Meaningful Artificial Intelligence: Towards a French and European Strategy", Cédric Villani et al., accessed May 7, 2019, https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf
[57] *See*, Yavar Bathaee, *"The Artificial Intelligence Black Box and The Failure of Intent and Causation"* Harvard Journal of Law & Technology 31,2 (2018): 890-938.

Chapter II – Legislative procedure in the European Union[58]

Before we can start properly thinking and discussing regulating AI at a European level it is paramount to understand how laws are made in the European Union. In other words, we need to understand the EU's legislative procedure. The Treaty of Lisbon changed substantially the law-making procedure within the EU and what type of legislative acts the EU can enact. Nowadays, if we are talking strictly about "hard law" as in binding legal act, the EU can enact either Regulations, Directives or Decisions. Decisions are not an adequate legal tool to establish the basis for a European-wide legal framework on AI[59]. Regulations are meant to completely harmonize the Member States legal framework, they are directly applicable and do not need implementation by national Act. The degree of harmonization pursued by a Directive can greatly vary depending on the subject and aims of the Directive. Maximum harmonization Directives leave few if any opportunities for "creative" transposition by Member States. However, there are aspects that will never be perfectly uniform when using Directives, including time of implementation (within the legal limitations, Member States will implement the necessary legal acts at their own description and according to their own scheduling)[60]. Directives are not directly applicable – even if they can produce direct effects under the right circumstances and if the criteria from established by the ECJ are met, namely being clear, unconditional and sufficiently precise[61] – and depend on Member States to implement them.

There are numerous factors that influence the choice between enacting a Directive or a Regulation. Competence is the first, in areas where the EU has exclusive or shared competence Regulation and Directives are both options at the legislator's disposal, but if

[58] This Chapter is heavily inspired by our previous work "Democracia, legitimidade e competência legislative na União Europeia", albeit with some relevant developments including the addition of trilogues and a schematic view of the ordinary legislative procedure. A general reference to the abovementioned work should be considered as existing through the entire chapter though. *See*, Tiago Sérgio Cabral, "Democracia, legitimidade e competência legislativa na União Europeia", *in* E-book UNIO/CONPEDI Vol. 2: Interconstitucionalidade: Democracia e Cidadania de Direitos na Sociedade Mundial - Atualização e Perspectivas, Alessandra Silveira coord. (Braga: CEDU, 2018), 265-292.

[59] Even though Article 288 TFEU allows for Decisions with a more general scope, and they have been used in matters as providing legal foundation for programmes, their most common use is when directed to a specific actor. Nevertheless, their constitutional design would never be adequate for the kind of endeavour that is building an uniformized legal framework.

[60] Generally, if the objective is to fully harmonize Member State law, we find Regulations to be the superior solution. They avoid problems such as poor or plain incorrect transposition, different transposition schedules, partial transposition, transposition dispersed through various national laws and minimise translation issues. They also offer more security to economic operators and citizens who just need to consult a specific legal act to know the applicable law, instead of searching through national legislation.

[61] *See*, Judgement of the ECJ of 5 April 1979, *Ratti*, Case C-148/78, ECLI:EU:C:1979:110.

the EU only has the competence to support, coordinate or supplement Member-State actions, opting for a Regulation would be in breach of the EU's constitutional law[62] (the flexibility clause contained in Article 352 could arguably be used, though). Political and social climate is the second, both regarding the EU and Member States. One can easily see how federalists or "*euroenthusiastics*" would be usually more open to supporting a Regulation than *eurosceptics*. A higher degree of satisfaction with the EU by its citizens can also allow for more intervention. A third factor is the nature of the Institutions responsible for producing laws in the EU. As *pure* European institutions both the Commission and the European Parliament are generally more ambitious than the Council whose approach is far more cautious.

Nevertheless, for both Directives and Regulations (and also for Decisions that are legal acts), there are only three procedures than can be followed: *i)* the ordinary legislative procedure (co-decision); *ii)* consent procedure and; *iii)* consultation procedure. The rules are as follows:

§ 1. Ordinary Legislative Procedure

Under the Treaty of Lisbon, the co-decision procedure was renamed, or one could argue rebranded has the ordinary legislative procedure. The principles and rules underlying it remain the same, but it should be now seen as the default legislative procedure and not on equal ground with the consent and consultation procedures. Its intended use was also vastly expanded. It is in this type of procedure that the EU's constitutional design draws closer to true bicameralism where both the Council and the EP legislate as true equals. The Parliament is the lower chamber and represents the European Citizens that directly elect MEPs while the Council is the higher chamber representing Member States's interests.

The right of initiative rests, with minor exceptions[63], with the European Commission. This means that neither the Council nor the European Parliament can directly propose new legislation, albeit both may request the EC to do so[64]. In fact, the

[62] In fact, this would also be the case for a Directive that intended to establish a maximum harmonization regimen.
[63] As an example, Article 76 TFUE permits exceptional initiative by a ¼ of the Member States, for Judicial Cooperation in Criminal Matters, Police Cooperation and for the creation of administrative measures to ensure cooperation within the Area of Freedom, Security and Justice.
[64] For the European Parliament this right is contained in Article 225 TFEU and for the Council in Article 241 TFEU.

Framework Agreement on relations between the European Parliament and the European Commission (hereinafter, "Framework Agreement") goes considerably further than the letter of the Treaties burdening the European Commission with the duty to *"to report on the concrete follow-up of any request to submit a proposal pursuant to Article 225 TFEU (legislative initiative report) within 3 months following adoption of the corresponding resolution in plenary. The Commission shall come forward with a legislative proposal at the latest after 1 year or shall include the proposal in its next year's Work Programme. If the Commission does not submit a proposal, it shall give Parliament detailed explanations of the reasons"*[65]. Some developments may happen soon on this front. In her opening statement in the European Parliament Plenary Session, the new President of the European Commission Ursula von der Leyen promised that, if elected[66], she would support a right of initiative for the European Parliament. It is still not clear how said support would materialize, it may be through trying to push an amendment to the Constitutional Treaties, even though that seems unlikely, or through a new Interinstitutional Agreement adopted under Article 295 TFEU. For now, the President-elect of the European Commission only specified that when Parliament adopts Resolutions requesting the Commission to submit legislative proposal, her Commission pledges to respond with a *"legislative act in full respect of the proportionality, subsidiarity, and better law-making principles"*. It is also unclear whether said legislative Resolution would contain a proposal for the content and text of the legislative act, that the EC would just "copy and paste" or if it just a request and content will still be defined by the EC. The first solution would be more akin to a true (albeit indirect) right to initiate legislation, the second would be similar to what is currently in the Framework Agreement. In this last scenario, the difference could be that the Commission would renounce the "right" to not submit a proposal[67].

The content of the EC's proposal has binding power over other actors in the legislative procedure. Amendments cannot completely change the nature of the proposal because that would be akin to undermining the power given to this Institution[68]. The

---

[65] Regarding the Council, Paul Craig and Gráinne de Búrca argue that this power is used with such frequency that in practice, it works almost like a pseudo-right of initiative. *See*, Paul Craig and Gráinne de Búrca, EU Law: Text, Cases and Materials, (6.ª ed., Oxford: Oxford University Press, 2015),125ff.

[66] The speech was delivered before the confirmation vote which von der Leyen won with a (thin) majority of 383 MEPs. The support of several MEPs from groups like S&D and Renew Europe was made conditional on including in her programme a number of proposals, including the one abovementioned.

[67] *See*, "Opening Statement in the European Parliament Plenary Session by Ursula von der Leyen, Candidate for President of the European Commission", European Commission, accessed July 17, 2019, http://europa.eu/rapid/press-release_SPEECH-19-4230_en.htm.

[68] Cfr. According to the Advocate General Tesauro *"the amendments adopted [cannot[fall outside the scope of the measure in question, as defined by the proposal"*. *See*, Judgment of the ECJ of 11 November 1997, *Eurotunnel SA v. SeaFrance*, Case C-408/95, ECLI:EU:C:1997:532, 35-39.; Opinion of the Advocate General delivered on 27

position of the EC is further strengthened by the fact that any amendment on which the Commission has delivered a negative opinion has to be approved with unanimity in the Council. The Commission can also alter its proposal at any time, as long as the Council has not yet acted.

The procedure's first reading is kickstarted by the submission of the EC's proposal to both the European Parliament and the Council. The Parliament is the first one to give its opinion and can choose three options: fully agree with it, amend it and send to the Council or reject it. If the European Parliament decides to reject the proposal at this stage the procedure ends, and the act is not adopted.

If the Parliament agrees or amends the proposal it moves on to the Council. The Council can vote by qualified majority to adopt the proposal in the wording which corresponds to the position of the European Parliament[69] or can adopt is own position on the proposed Act and then send it back to the Parliament.

If the co-legislators are not able to agree on a position in the first reading, the procedure enters into its second reading. The proposal goes back to the European Parliament where it can be approved by a majority of the votes cast or by tacit agreement (absence of a decision within a three-month time limit, extendable by one month). The proposal can also be rejected or amended by a majority "of its component members". Rejection by Parliament means the end of the line for the proposal. However, Parliamentary amendments mean that the proposed Act must go back to the Council where, by qualified majority, the Institution must choose one of the following; either *i)* accept all amendments or *ii)* reject all amendments. There is no more room for back-and-forth modifications between co-legislators. If the Council fails to accept the Parliament's demands, a Conciliation Committee is created.

Conciliation aims to bring the legislative procedure to a successful conclusion through closer cooperation between Parliament and Council, fostered by the Commission[70]. The Committee is composed of an equal number of representatives from the Council[71] and from the European Parliament whose duty is to reach a joint proposal

May 1997, *Eurotunnel SA v. SeaFrance*, Case C-408/95, ECLI:EU:C:2014:2470; Judgment of the ECJ of 14 April 2015, *Council v Commission*, Case C-409/13, ECLI:EU:C:2015:217.

[69] If no amendments are made to the proposal by the European Parliament, it can directly reflect the wording proposed by the European Commission.

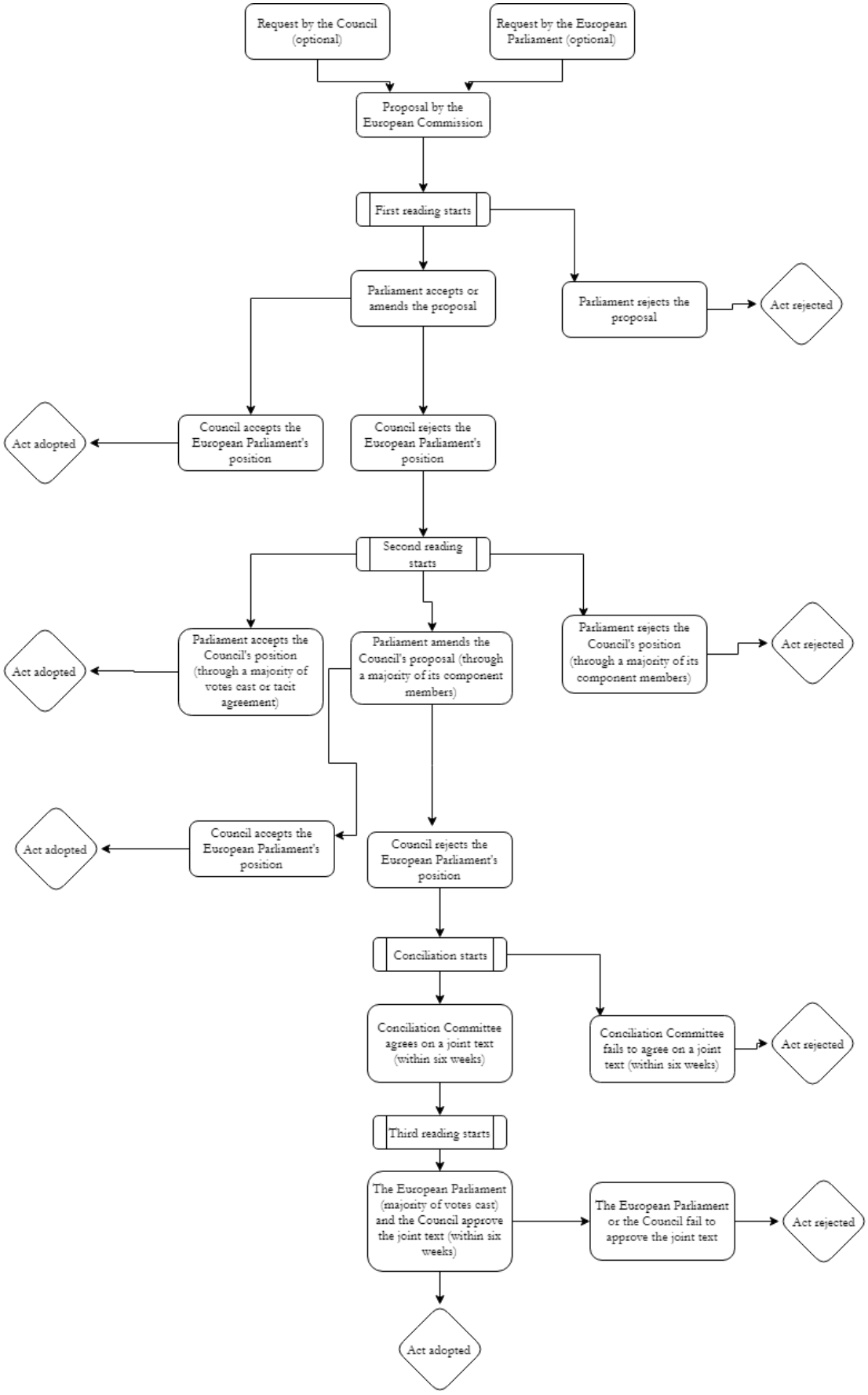
[70] One must not that trilogues can also take place before and during Conciliation to increase the likelihood of a fruitful conclusion.

[71] Representatives from the Council can be either members of the Council or their representatives

(Article 294(10) TFEU). The Committee must reach this joint proposal within a 6-week timeframe, otherwise the proposal is not adopted.

If the Conciliation Committee reaches a joint proposal, the third, and last, reading begins. Both Council and Parliament have a 6-week deadline to approve the joint agreement. Parliament acts by majority of the votes cast and the Council acts by qualified majority. This is the final opportunity for the co-legislators to adopt the act, if it is not approved, the legislative procedure ends here[72].

[72] *See*, Miguel Gorjão-Henríques, "Anotação ao art.º 289.º do TFUE", in, *Tratado de Lisboa anotado e comentado,* coord. Manuel Lopes Porto and Gonçalo Anastácio, (Coimbra: Almedina, 2012), 1034-1038; Inês Morgado, "Anotação ao art.º 293.º do TFUE", in *Tratado de Lisboa*…, 1048-1049; César Cortes e Paulo Rangel, "Anotação ao art.º 294.º do TFUE", in *Tratado de Lisboa*…, 1050-1059; Robert Schiitze, *European constitutional law*, (Cambridge: Cambridge University Press, 2012), 169ff.; João Mota de Campos, João Luís Mota de Campos e António Pinto Ferreira, *Manual de direito europeu: o sistema institucional. A ordem jurídica e o ordenamento económico da União Europeia*, (7th ed., Coimbra: Coimbra Editora, 2014), 214ff.; Alina Kaczorowska, *European Union Law*, (3.ª ed., London: Routledge, 2013); 148ff.; Directorate-General for Internal Policies of the Union and Directorate for Legislative Coordination and Conciliations Conciliations and Codecision Unit, *Handbook on the Ordinary Legislative Procedure* (European Parliament: Brussels, 2017), 11-25 and 37-39; Pierre Mathijsen and Peter Dyrberg, *Mathijsen's Guide to European Union Law* (11th ed., London: Sweet & Maxwell Ltd, 2013). 88ff; Vincenzo Guizzi, *manuale di diritto e politica dell' Unione Europea* (4th ed., Naples: Editoriale Scientifica: 2015), 366ff; Miguel Gorjão-Henriques, *Direito da União: História, Direito, Cidadania, Mercado Interno e Concorrênncia* (9th ed., Coimbra: Almedina, 2019), 249ff.


Request by the Council (optional)
Request by the European Parliament (optional)
Proposal by the European Commission
First reading starts
Parliament accepts or amends the proposal
Parliament rejects the proposal
Act rejected
Act adopted
Council accepts the European Parliament's position
Council rejects the European Parliament's position
Second reading starts
Act adopted
Parliament accepts the Council's position (through a majority of votes cast or tacit agreement)
Parliament amends the Council's proposal (through a majority of its component members)
Parliament rejects the Council's position (through a majority of its component members)
Act rejected
Act adopted
Council accepts the European Parliament's position
Council rejects the European Parliament's position
Conciliation starts
Conciliation Committee agrees on a joint text (within six weeks)
Conciliation Committee fails to agree on a joint text (within six weeks)
Act rejected
Third reading starts
The European Parliament (majority of votes cast) and the Council approve the joint text (within six weeks)
The European Parliament or the Council fail to approve the joint text
Act rejected
Act adopted

OLP: Diagram

§ 1.1. Trilogues

Currently, 97% of all OLP files are adopted either at first or early second reading, with the percentage of full second readings standing at a mere 3% and third readings having practically disappeared. The reason for this phenomenon is the growing trend of trilogues. trilogues are informal (but institutionalized) *"negotiations between representatives of Parliament and Council, assisted by the Commission, aimed at reaching agreement on legislation, normally at an early stage of the legislative process"*[73/74].

Typically, a variable number of trilogue meetings is held between representatives of the three law-making institutions, about 3.5 being the average[75]. Generally, trilogues occur in the Parliament's Brussels building. Technical trilogues like preparatory or operational meetings can be used to clear the path for the main political trilogues.

In the trilogue meetings, *"Parliament is represented the Committee Chair leading the delegation, sometimes replaced by a Vice-President, the Rapporteur and Shadow Rapporteurs from the different parties, their assistants, political party functionaries and Committee secretariat staff. The Council is represented at earlier stages by the civil servant who chaired the Council Working Party supported by the Council Secretariat, and at later stages by the Chair of the Committee of Permanent Representatives. Finally, the Commission sends Heads of Unit or Directors, supported by the Legal Service and the Co-Decision Unit, although sometimes Director-Generals or their Deputies attend from the outset. A Commissioner often attends the concluding trilogues"*[76].

Negotiations are based on a multi-column document where each of the first three columns contains the positions of the Commission, Parliament and Council. The fourth column is initially blank and is filled during negotiation with the jointly agreed position.

---

[73] *See*, "Decision of the European Ombudsman setting out proposals following her strategic inquiry OI/8/2015/JAS concerning the transparency of Trilogues", European Ombudsman, access January, 20, 2019, https://www.ombudsman.europa.eu/en/decision/en/69206; Gijs Jan Brandsma, *"Transparency of EU informal trilogues through public feedback in the European Parliament: promise unfulfilled",* Journal of European Public Policy (2018): 1-20.

[74] Currently 70% to 80% of EU legislation is adopted following trilogues. See, Judgment of the EGC of 22 March 2018, *De Capitani v Parliament*, Case T-540/15, ECLI:EU:T:2018:167, 70.

[75] More complex legislation might take significantly more trilogues to be agreed. The General Data Protection Regulation, for example, took 14 trilogues before the law-making institutions could reach an agreement.

[76] *See,* Justin Greenwood and Christilla Roederer-Rynning, "Taming Trilogues: The EU's Law-Making Process in a Comparative Perspective", in *The European Parliament in times of EU crisis: Dynamics and Transformations*, Olivier Costa ed., (Maastricht: Palgrave Macmillan, 2019), 121-141; Deirdre Curtin & Päivi Leino, *"In Search of Transparency for EU Law-Making: Trialogues on the Cusp of Dawn",* Common Market Law Review 54,6 (2017): 1673–1712

The three institutions hail trilogues as highly useful tools that contribute to expedient law-making. And while it is undeniable that they are effective, we cannot ignore the fact that Authors have been, for a number of years, pointing that when compared to the constitutionally enshrined readings procedure trilogues lack in transparency.

Trilogues are closed-door affairs attended by a few representatives of the three Institutions[77]. Documents related to trilogues are not freely available to citizens and, in fact, the Institutions usually fight tooth and nail against attempts to access them. They argue that making the discussions public causes disturbances in the legislative procedure and does allow for the legislators to freely express their opinion[78]. We deem such a position to unacceptable for both co-legislators, but especially for the European Parliament.

Regarding the Council, being non-transparent is its *modus operandis*. Council meetings are relatively informal, ministers even address each other on a first-name basis. This ensures that, away from the public eye, they can negotiate freely and compromise without fearing backlash in their Member States. The Council prefers to see its reunions as diplomatic summits instead of what they really are: legislative meetings from a Senate-like Chamber of a law-making body. Council documents regarding legislative files are usually only made available to the public after the procedure is over. Moreover, some documents like Legal Service opinions are generally not disclosed, in this case as to not restrict the freedom of the legal counsel. Member States can also request that documents reflecting their position during negotiations not be made available to the public.

We should not forget that when legislating accountability is key, citizens should be able to analyse and, if they disapprove of the way that the minister is conducting business in the Council, vote his government out. However, truth be told, the Council is not

---

[77] A frequent criticism to trilogues is that this type of procedure puts too much power in a few representatives/officials. However, empirical analysis seems to contradict these conclusions. *See*, Anne Rasmussen & Christine Ren, *"The consequences of concluding codecision early: trilogues and intra-institutional bargaining success"* Journal of European Public Policy 20,7 (2013): 1006-1024.

[78] *See*, Directorate-General for Internal Policies of the Union and Directorate for Legislative Coordination and Conciliations Conciliations and Codecision Unit, *Handbook*…., 26-36.; Christilla Roederer-Rynning, *"Passage to bicameralism: Lisbon's ordinary legislative procedure at ten"*, Comparative European Politics (2018): 1-17; Directorate-General for Internal Policies of the Union and Directorate for Legislative Coordination and Conciliations Conciliations and Codecision Unit, *Activity Report on the Ordinary Legislative Procedure: 4 July 2014 - 31 December 2016 (8th parliamentary term)* (European Parliament, Brussels: 2017), 19ff.; Christilla Roederer-Rynning and Justin Greenwood, *"The culture of trilogues"*, Journal of European Public Policy 22,8 (2015): 1148–1165; Christilla Roederer-Rynning and Justin Greenwood, *"The European Parliament as a developing legislature: coming of age in trilogues?"*Journal of European Public Policy 24,5 (2017): 735-754.
"How transparency can be improved in the way EU laws are negotiated and agreed", Aidan O'Sullivan, access January 30, 2019, https://blogs.lse.ac.uk/europpblog/2016/08/18/how-transparency-can-be-improved-in-the-way-eu-laws-are-negotiated-and-agreed/;

designed to be and as never tried to assert itself as the Institution that, more closely, represents the will of the European citizens. That is the role of the EP.

Indeed, MEPs are elected to the European Parliament by direct universal suffrage (Article 14(3), TEU) and have the most direct democratic legitimacy of any of the Institutions[79]. Parliament takes this to heart, and most legislative affairs there happen in a fairly transparent manner. Both plenary and committee meetings are broadcasted live and made available with interpretation to allow citizens from different countries to understand them. Getting accreditation to access the European Parliament is a relatively straightforward procedure. Furthermore, lobbying is regulated and done with increasing transparency[80]. With all this in mind, are there any efforts by the European Parliament to boost transparency in trilogues? The short, answer is that yes there are, but they do not seem to be enough. In a 2014 Resolution on public access to documents the European Parliament argued for broad changes to trilogues by calling *"on the Commission, the Council and Parliament to ensure the greater transparency of informal trilogues, by holding the meetings in public, publishing documentation including calendars, agendas, minutes, documents examined, amendments, decisions taken, information on Member State delegations and their positions and minutes, in a standardised and easy accessible online environment[81]"*. More recent Resolutions have restated Parliament's position on the issue, even though the demand for public meetings has apparently disappeared[82].

The other Institutions have shown little receptivity to these proposals and, in fact, even Parliament did little of substance to achieve its aims. Currently, Article 69, point f) of the EP's Rules of Procedure enshrines rules about the mandatory composition of negotiation teams to avoid any excessive concentration of power in a reduced number of

---

[79] *See*, Tiago Sérgio Cabral and Rita de Sousa Costa, *"The European Union's existential crisis: current challenges from populism to Donald Trump"* UNIO - EU Law Journal.4,1 (2018): p. 3-15

[80] In January 2019, the European Parliament approved an amendment to Rules of Procedure that orders *"rapporteurs, shadow rapporteurs and committee chairs shall, for each report, [to] publish online all scheduled meetings with interest representatives falling under the scope of the Transparency register"*. Negotiations between the Institutions are underway to revamp and expand the Transparency Register.

[81] *See*, "European Parliament resolution of 11 March 2014 on public access to documents (Rule 104(7)) for the years 2011-2013", European Parliament, access January 20, 2019, https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0203+0+DOC+XML+V0//EN&language=EN.

[82] *See*, "European Parliament resolution of 14 September 2017 on transparency, accountability and integrity in the EU institutions", European Parliament, access January 23, 2019, https://www.europarl.europa.eu/sides/getDoc.do?type=TA&language=EN&reference=P8-TA-2017-0358; "European Parliament resolution of 28 April 2016 on public access to documents (Rule 116(7)) for the years 2014-2015", European Parliament, access January 23, 2019, https://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P8-TA-2016-0202&language=EN#def_1_10.

actors (Article 69. point f), first paragraph). There also rules on "reporting back" the progress achieved during trilogues (Article 69, point f), third paragraph). The underlying reasoning behind reporting back rules is that if the negotiators report to the responsible committees and since said committees are broadcasted to the public, citizens can also get the relevant information. In practice, the rule does not apply if it is not "feasible" to do so, and studies show that it plainly does not work[83].

The problem is considered severe enough to merit the European Ombudsman's attention. In her Decision setting out proposals following her strategic inquiry OI/8/2015/JAS concerning the transparency of trilogues, the Ombudsman states the current structure of trilogues might breach the principles of representative democracy (Article 10 TUE) and the right of European citizens to participate in the democratic life of the Union. The issue is, above all else, one of accountability. If citizens and national parliaments do not know how their representatives cast their votes, how they are making their arguments for the causes they represent when they are making concessions, and when they are standing their ground, they cannot properly assess their performance and penalize or reward them accordingly. Furthermore, it is close to impossible for citizens and stakeholders to participate in the decision-making if they do not know when the meetings are being held, who is there, what is being discussed and is the current stage of the discussions. Lobbying is an integral part of the European Union, one might even say a cornerstone, but lobbying should not be understood as practice that is only achievable by a select few insiders. Every citizen should be able to lobby their representatives in a transparent manner.

However, it is a fact that public pressure and continuous vigilance over the representatives' behaviour might be counterproductive, especially in the middle of the negotiation procedure. A representative might lose the ability to think strategically and to give ground on some issues expecting to gain concessions further down the line because public opinion reacts immediately, and public outrage spreads like wildfire.

After weighing all these considerations, the Ombudsman made the following proposals:

[83] Furthermore, the voting rules in Article 69 take negotiating power away from the Parliament by not allowing amendments to the provisional agreement. *See*, Justin Greenwood and Christilla Roederer-Rynning, "Taming Trilogues: The EU's Law-Making Process…", 127-130.

1. *That the institutions make publicly available a "trilogue calendar" identifying forthcoming trilogues. They should also refer to trilogues in databases on legislative files;*
2. *That both co-legislators make proactively available, before trilogue negotiations begin, their positions on the Commission proposal, regardless of the level at which the position has been adopted internally and regardless of the legislative proposal;*
3. *That the Institutions make available general summary agendas before or shortly after trilogue meetings;*
4. *That the institutions make proactively available four-column documents, including the final agreed text, as soon as possible after the negotiations have been concluded;*
5. *That the institutions include, in legislative databases and calendars dealing with trilogues, links to any minutes or videos of the institutions' public meetings where a trilogue has been discussed;*
6. *That the institutions make proactively available a list of the representatives who are politically responsible for decisions taken during a trilogue, such as the MEPs involved, the responsible Minister from the Council Presidency and the Commissioner in charge of the file. If the power to take decisions is delegated to civil servants, their identities should also be disclosed proactively;*
7. *For the purposes of facilitating requests for public access to documents, the institutions make available as far as possible lists of documents tabled during trilogue negotiations;*
8. *Furthermore, the Ombudsman encourages the institutions to work together to make as much trilogue information and documentation as possible publicly available through an easy-to-use and easy-to-understand joint database*[84].

The Ombudsman's proposals are, in our opinion, generally balanced and easy to apply. If anything, one could argue that they should go further on issues like access to the multi-column documents. In this matter the Ombudsman seems to accept a type of presumption of secrecy for the document until their content can no longer have an impact on negotiations, going opposite to the European Court of Justice and General Court's case-law on this matter[85]. The general rule should be for open access to documents related

[84] "Decision of the European Ombudsman setting out proposals following her strategic inquiry OI/8/2015/JAS concerning the transparency of Trilogues", European Ombudsman, accessed January 20, 2019, https://www.ombudsman.europa.eu/en/decision/en/69206

[85] Fact is that the European Court of Justice's interpretation of what can seriously undermine the decision-making is extremely strict as seen in *Access Info Europe*. *See*, Judgment of the ECJ of 17 October 2013, *Access Info Europe v. Council*, Case C-280/11 P, ECLI:EU:C:2013:671, upholding the Judgment of the EGC of 22 March 2011, Case T-233/09, ECLI:EU:T:2011:105.

to the legislative procedure, with exceptions only tolerable when disclosure of the document would seriously undermine the institution's decision-making process. Such an issue should be decided on a case-by-case basis. We can accept that general presumptions of confidentiality have been recognized by the European Court Justice for some types of documents, namely: *"the documents in an administrative file relating to a procedure for reviewing State aid; the submissions lodged in proceedings before the courts of the European Union, for as long as those proceedings remain pending; the documents exchanged between the Commission and notifying parties or third parties in the course of merger control proceedings; the documents relating to an infringement procedure during its pre-litigation stage, including the documents exchanged between the Commission and the Member State concerned during an EU Pilot procedure; and the documents relating to a proceeding under Article 101 TFEU*"[86]. However, there was no reason to think that these general presumptions of confidentiality also applied to every multi-column document. In fact, after *De Capitani* it seems clear that multi-column documents in general are, indeed, not protected by this exception and should be disclosed as widely as possible, as a general principle. Full refusal should be a last resort, even when the requirements under Article 4 of Regulation 1049/2001/EC are met, solutions like confidentially obligations to be imposed on the interested parties should be studied before opting for a negative answer.

In *De Capitani*[87], the General Court of the European Union analysed the refusal by the European Parliament to provide access to the fourth column of two multi-column documents part of trilogue negotiations on Police Cooperation, "*including issues of data protection and the management board of"* Europol. The European Parliament sustained its refusal on Article 4(3) of Regulation nº 1049/2001/EC of the European Parliament and of the Council of 30 May 2001 regarding public access to European Parliament, Council and Commission documents arguing that making these documents available would impair the Institutions' ability decide, harm the confidence between the Institutions and Member States, hinder cooperation with the Council and open to door to external pressures on the ongoing negotiation.

We note that, with more than a touch of incoherence, Parliament appears in contradiction with its previous Resolutions. Even though the Resolutions accept the existence of exceptions to the principle of wide access by citizens to the legislative

---

[86] *See*, Judgment of the ECJ of 4 September 2018, *ClientEarth v Commission*, Case C-57/16 P, ECLI:EU:C:2018:660 and also the decision that it repealed, Judgment of the EGC of 13 November 2015, *ClientEarth v Commission*, Case T-424/14, ECLI:EU:T:2015:848.

[87] See, Judgment of the EGC of 22 March 2018, *De Capitani v Parliament*, Case T-540/15, ECLI:EU:T:2018:167.

procedure, their spirit is that exceptions are precisely that: exceptions. Such a broad interpretation of the exceptions under Regulation nº 1049/2001/EC clearly goes against their spirit (and also of the 2016 interinstitutional agreement).

With all this in mind, the Court's ruling can only be seen as a victory to transparency and a reason for the Institutions to rethink their position on the issue. First, the Court states that the exceptions in Article 4 of Regulation nº 1049/2001/EC must be interpreted and applied strictly[88]. Furthermore, the fact that a matter is covered by one of the exceptions is not enough to warrant refusal, a casuistic analysis of the particular case must be conducted. Only if the danger is likely to manifest itself and that likelihood of manifestation is reasonably foreseeable can the exception be called upon.

The Court also clarified in definitive that trilogues are part of the legislative procedure and thus subject to all general principles that govern it, including openness and transparency. Furthermore, the Court found that there is no general presumption of non-disclosure regarding the fourth column of the multi-column document.

Regarding accountability and the right to participate in the democratic life of the Union, the Court argued that the European citizens are perfectly capable of understanding the nature of trilogues and the fact that the positions expressed in the fourth column can change during negotiation or even end with no deal at all. The paternalistic vision of the Institutions seems to have been rejected. Additionally, the General Court delivered a much-needed reminder that cooperation between the Institutions is not a choice at their discretion but a constitutional obligation.

Following this reasoning, the Court decided to annul the decision refusing access to the multi-column document. The Committee on Legal Affairs of the European Parliament decided to accept the decision instead of submitting an appeal to the European

[88] In its Judgment in *Turco* the European Court of Justice shows how strict its interpretation is. Even if a document is covered by one of the exceptions of Regulation nº 1049/2001/EC, in this case regarding legal opinions, a comprehensive assessment must be undertaken to decide which parts, in particular, cannot be disclosed. *See*, Judgment of the ECJ of 1 July 2008, *Turco v. Council*, Joined Cases C-39/05 P and C-52/05 P, ECLI:EU:C:2008:374.

Court of Justice. However, efforts to reinforce transparency in its wake have, thus far, been meek[89/90].

Questions can, probably will, and should indeed arise around the question of access and democratic legitimacy of legislation relating to AI if citizens are not able to follow the negotiation between the various institutions, make their opinion known and influence it. This may boost any mistrust concerning AI harbored by the public at large but, more than that, heavily opens the door for a highly damaging, and possibly effective, populist messages: "the bureaucrats in Brussels are creating rules governing a potentially world ending technology in dark damp rooms". Obviously, feeding into such rhetoric should be avoided at all costs.

§ 2. Special Legislative Procedures

§ 2.1. Consent Procedure

Following the general rule, in the consent procedure, the right to propose legislation lies with the EC[91]. Still, unlike in the ordinary legislative procedure, within the limits of the ECJ's case-law (namely, amendments cannot completely change the nature of the proposal, as stated above) the Council possesses full decision power on the legal act's content. In this type of procedure, the EP maintains vetoing power but no more. It can either "agree" with the act in its entirety or stop it completely. In practice, this "take it or leave it" approach is softened by the EP's power to issue an interim report during the procedure and by informal discussions between the Institutions. The EP may use these tools to suggest the necessary amendments to win parliamentary approval[92]. The consent procedure is used in situations such as:

---

[89] *See*, Francesca Martines, *"Transparency of Legislative Procedures and Access to Acts of Trilogues: Case T-540/15, De Capitani v. European Parliament"* European Papers 3,2 (2018):947-959; EU "Court Condemns the EU Legislative Process for Lack of Transparency: Time to Open Up?", Massimo Frigo, accessed December 10, 2018, http://opiniojuris.org/2018/03/27/33507/.

[90] "Interinstitutional Agreement between the European Parliament, the Council of the European Union and the European Commission on Better Law-Making", EC, EP and Council, accessed December 10, 2018, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016Q0512(01)&from=EN

[91] An exception can be identified in Article 86(1) TFEU (regarding the implementation of the European Public Prosecutor's Office) read along with Article 76 TFEU.

[92] Alina Kaczorowska, *European*…, 158 e ss; Paul Craig and Gráinne de Búrca, *EU Law*…, 130 e ss; Robert Schiitze, *European Constitutional*…, 176 e ss; "Guide to EU decision-making and justice and home affairs after the Treaty of Lisbon", Steve Peers, accessed December 13, 2019, http://www.statewatch.org/analyses/no-115-lisbon-treaty-decision-making.pdf; "The Special Legislative Procedures: Consent", University of Portsmouth European Studies Hub, accessed December 13, 2019,

- Article 7 (1) TEU, according to which on a reasoned proposal by one third of the Member States, by the European Parliament or by the European Commission, the Council, acting by a majority of four-fifths of its members after obtaining the consent of the European Parliament, may determine that there is a clear risk of a serious breach by a Member State of the values referred to in Article 2 TEU.

- Article 7 (2) TEU, which establishes that the European Council, acting by unanimity on a proposal by one third of the Member States or by the Commission and after obtaining the consent of the European Parliament, may determine the existence of a serious and persistent breach by a Member State of the values referred to in Article 2 TEU, after inviting the Member State in question to submit its observations[93].

- Article 49 TUE, regarding application and accession to the Union states that the Council shall act unanimously after consulting the Commission and after receiving the consent of the European Parliament, which shall act by a majority of its component members and respecting the conditions of eligibility agreed upon by the European Council[94].

- Article 50 TEU, regarding the agreement setting out the arrangements for the withdrawal of a Member State from the Union. The abovementioned agreement shall be concluded on behalf of the Union by the Council, acting by a qualified majority, after obtaining the consent of the European Parliament[95].

- Article 25 TFEU, regarding the reinforcement or addition of rights to the European citizen. In this situation the Council acting unanimously in accordance with a special legislative procedure and after obtaining the consent of the

http://hum.port.ac.uk/europeanstudieshub/learning/module-2-understanding-eu-policy-making/the-special-legislative-procedures/.

93 *See*, Stelio Mangiameli and Katharina Pabel, "Article 7 [The Principles of the Federal Coercion]", in *The Treaty on European Union (TEU): a Commentary,* Hermann-Josef Blanke and Stelio Mangiameli eds., (Heidelberg: Springer, 2013), 349-373; Luís Miguel País Antunes, , "Anotação ao artigo 7.º do TUE", in *Tratado de Lisboa*…, 43-46.

94 *See*, Susanna Fortunato, "Article 49 [Accession to the Union]", in *The Treaty on European Union (TEU)*…, 1357-1383; Ricardo Bayão Horta, "Anotação ao artigo 49.º do TUE", in *Tratado de Lisboa*…, 183-185.

95 *See*, note 146

European Parliament may adopt provisions with this aim. Said provisions will only enter into force after approval by Member States[96].

- Article 218 (6) TFEU, when the Council adopts a decision concluding the following agreements it must obtain the consent of the EP: *i)* association agreements; *ii)* agreement on Union accession to the European Convention for the Protection of Human Rights and Fundamental Freedoms; *iii)* agreements establishing a specific institutional framework by organising cooperation procedures; *iv)* agreements with important budgetary implications for the Union; *v)* agreements covering fields to which either the ordinary legislative procedure applies, or the special legislative procedure where consent by the European Parliament is required[97].

- Article 223 (1) TFEU, covering the election rules for the European Parliament[98].

- Article 311 (4th paragraph) TFEU, according to which the Council, acting by means of regulations in accordance with a special legislative procedure, shall lay down implementing measures for the Union's own resources system in so far as this is provided for in the decision adopted on the basis of the third paragraph. The Council shall act after obtaining the consent of the European Parliament[99].

- Article 352 TFUE (flexibility clause)[100].

Infrequently, there are some residual cases where the European Parliament acts as the commanding institutions. Those coincide with the (very few) situations where the EP

---

96 *See*, Rui Manuel Moura Ramos, "Anotação aos art.os 18.º a 23.º do TFUE", in *Tratado de Lisboa*…, 258-263; Rui Manuel Moura Ramos, "Anotação aos artº 25 do TFUE", in *Tratado de Lisboa*…, 268:

97 *See,* Margarida Afonso, "Anotação ao art.º 218.º do TFUE", in *Tratado de Lisboa*…, 832-837.

98 *See*, Vital Moreira, "Anotação ao art.º 223.º do TFUE", in *Tratado de Lisboa*…, 847-850.
While we concede that this procedure possesses very specific characteristics such as the such as the fact that the EP must draw up an initial proposal, this is not enough to make one perish to thought that this provision is in need of serious revamp. At the very least, the ordinary legislative procedure should have been adopted in this situation. It is unacceptable and, in fact, borderline embarrassing for the entirety of European democracy that the European Parliament has less power to decide the rules of its own election than the Council.

99 *See*, Manuel Lopes Porto, "Anotação ao art.º 311.º do TFUE", in *Tratado de Lisboa*…, 1099-1102.

100 *See*, Ana Maria Guerra Martins, "Anotação ao art.º 352.º do TFUE", in *Tratado de Lisboa*…, 1232-1235.

has legislative initiative and generally relate to matters of internal organisation. Some examples are Article 223 (2) TFEU, Article 226 TFEU[101] and Article 228 (4) TFEU.

§ 2.2. Consultation Procedure

In the consultation procedure, the Council appears in an even stronger position (in comparison with the consent procedure). Again, the right of initiative rests with the Commission and this Institution refers the proposal to the Council (constitutional limitations to the degree of changes that can be made to the initial proposal still apply)[102]. In this procedure, the Council is bound to request the opinion of the EP regarding the Act, but no more than that. The EP has no veto power in this procedure and the substantive content of its opinion holds no more than a persuasive nature over the Council. If the Council wishes to do so, it may ignore it its entirety and decide in an entirely different manner. Doing so does not affect the validity of the Act.

Nevertheless, Parliamentary opinion must be heard (even if not heeded) and the Council cannot suppress or ignore this step. It is an essential formality in accordance with the ECJ's judgment in *SA Roquette Frères v. Council*. In this judgment the Court declared Regulation 1293/79/CEE to be void due to not having complied with this requirement. According to the ECJ: *"the consultation provided for in the third subparagraph of Article 43 (2), as in other similar provisions of the Treaty, is the means which allows the Parliament to play an actual part in the legislative process of the Community. Such power represents an essential factor in the institutional balance intended by the Treaty. Although limited, it reflects at Community level the fundamental democratic principle that the peoples should take part in the exercise of power through the intermediary of a representative assembly. Due consultation of the Parliament in the cases provided for by the Treaty therefore constitutes an essential formality disregard of which means that the measure concerned is void"*[103]. This parliamentary power is not without limits though. According to the ECJ's in *Parliament v. Council*, inter-institutional dialogue is subject *"to the same mutual duties of sincere cooperation as those which govern relations between Member States and the Community institutions"*. Thus "*Parliament is not entitled to complain of the Council's failure to await its opinion before adopting*" an Act when

[101] Regarding the rules enshrined in Article 226 TFEU, the appropriateness of the EP having to request consent from the EC and Council to define rules governing the right of enquiry in the EP is debatable.
[102] In line with what we explained regarding the consent procedures, there are also some exceptions in the consultation procedures such as Articles 87(3) and 89 TFEU, read in connexion with Article 76 TFEU.
[103] *See*, Judgment of the ECJ of 29 October 1980, *SA Roquette Frères v. Conselho*, Case 138/79, ECLI:EU:C:1980:249, 33-36.

failure to comply with the *"essential procedural requirement of Parliamentary consultation"* resulted from *"Parliament's failure to discharge its obligation to cooperate sincerely with the Council"*[104].

Nonetheless, one may argue that the – admittedly very limited, especially if we take into account the limitations derived from the ECJ's case-law –power to slow down an Act's adoption may grand some influence to EP in a legislative procedure where it is wholly outranked and outgunned by the Council. By withholding its opinion, it is (at least theoretically) possible for Parliament to hold informal negotiations that may be particularly effective when the proposal is urgent in nature, when the EP's position finds wide support within the ranks of its MEPs, when the EP's position is in line with the Commission's and when the Act requires unanimity to be approved in the Council. If the proposed Act is part of a legislative package containing measures that follow other types of legislative procedure (consultation and especially the ordinary legislative procedure), Parliament may make its approval on other acts in the package depended on amendments to the Act(s) adopted under the consultation procedure[105/106].

---

104 *See*, Judgment of the ECJ of 16 July 1992, *Parliament v. Council*, Case C-65/93 ECLI:EU:C:1992:325, 23-27.

105 Raya Kardasheva, "*The Power to Delay: The European Parliament's Influence in the Consultation Procedure*", Journal of Common Market Studies 47,2 (March 2009): 385-409.

106 The power to delay is frequently used in other Constitutional Systems as a negotiation tool. In the United States of America, under the right circumstances Filibusters may might allow a majority to delay, negotiate on even stop the adoption of undesirable measures. "Shutdowns" may also be considered as a means to use time as a factor of pressure in negotiations. The talking filibuster is the most well-known, but not the only type of filibuster, in the talking filibuster a member of the law-making body speaks continuously for long hours, delaying the debate regarding a certain proposal and hindering its adoption. Legal scholars and political science authors are divided regarding the use of these and similar tools, some consider that their excessive use halts progress and does not respect the will of the majority. Others counterargue that these tools are essential in protection the rights of the minority. These positions are sometimes inconsistent and tend to vary in accordance with the political party who is in power. Regarding filibusters *see*, Gregory Koger, *Filibustering: A Political History of Obstruction in the House and Senate* (Chicago: The University of Chicago Press, 2010); Catherine Fisk and Erwin Chemerinsk, "*The Filibuster*", Stanford Law Review 49,2 (1997): 181-254; David R. Jones, "*Explaining restraint from filibustering in the US senate*", The Journal of Legislative Studies 6,4 (2000): 53-68; Patrick Fisher, "*The filibuster and the nature of representation in the United States Senate*", Parliaments, Estates and Representation 26,1 (2006):187-195; Martin B. Gold e Dimple Gupta, "*The Constitutional Option to Change Senate Rules and Procedures: A Majoritarian Means to Overcome The Filibuster*", Harvard Journal of Law and Public Policy 28,1 (2005): 205-272; Michael J. Gerhardt, "*The Constitutionality of the Filibuster*", Constitutional Commentary 21 (2005): 445-484; Ernest Bormann, "*The southern senators' filibuster on civil rights: Speechmaking as parliamentary stratagem*", The Southern Speech Journal 27,3 (1962): 183-194; Steven S. Smith e Hong Min Park, "*Americans' Attitudes About the Senate Filibuster*", American Politics Research 41,5 (2013):735-760; Sarah A. Binder, Eric D. Lawrence e Steven S. Smith, "*Tracking the Filibuster, 1917 to 1996*" American Politics Research 30,4 (2002): 406-422; Josh Chafet, "The Unconstitutionality of the Filibuster", Connecticut Law Review 43,4 (2011):1003-1040; Gerard N. Magliocca, "*Reforming the Filibuster*", Northwestern University Law Review 105,1 (2011): 303-328; Emmet J. Bondurant, "*The Senate Filibuster: The Politics Of Obstruction*", Harvard Journal on Legislation 48 (2011):467-514; Laura T. Gorjanc, "*The Solution to the Filibuster Problem: Putting the Advice Back in Advice and Consent*", Case Western Reserve Law Review 54,4 (2004): 1435-1463; Brent Wible, "*Filibuster vs. Supermajority Rule: From Polarization to a Consensus- and Moderation – Forcing Mechanism for Judicial Confirmations*", William & Mary Bill of Rights Journal 13,3 (2005): 923-965.
Regarding shutdowns, their motives and effects *see.* "Shutdown of the Federal Government: Causes, Processes, and Effects", Clinton T. Brass, accessed December 10, 2018,

Some notable examples of the use of the consultation procedure are:

- Article 21(3) TFEU, allowing Council – absent (specific) provision of the necessary powers by the Treaties – to adopt measures concerning social security or social protection to protect the right of Union citizens to move and reside freely within the territory of the Member States.

- Article 74 TFEU, regarding the adoption of measures to ensure administrative cooperation between Member States within the Union's area of freedom, security and justice[107].

- Articles 150 and 160 TFEU, regarding the establishment of the Employment Committee and the Social Protection Committee.

- Article 218(6) TFEU, when the consent procedure is not necessary.

- Article 311 (3rd paragraph) TFUE, regarding the adoption of provisions relating to the system of own resources of the Union. The Council's decision will not

http://digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=2182&context=key_workplace; "Shutdown of the Federal Government: Causes, Effects, and Process", Kevin R. Kosar, accessed December 11, 2018, http://assets.thefiscaltimes.com/TFT2_20101228/App_Data/MediaFiles/1/B/1/%7B1B124168-264B-4686-8E6F-DE0D5F0E097E%7DShutdown%20background.pdf; Roy T. Meyers, *"Late Appropriations and Government Shutdowns: Frequency, Causes, Consequences, and Remedies"* Public Budgeting & Finance 17,3 (1997): 25-38; Scott R. Baker e Constantine Yannelis, *"Income Changes and Consumption: Evidence from the 2013 Federal Government Shutdown"* Review of Economic Dynamics 23 (2017): 99-124; Corinne Bendersky, *"Resolving ideological conflicts by affirming opponents' status: The Tea Party, Obamacare and the 2013 government shutdown"* Journal of Experimental Social Psychology 53 (2014): 163-168; "Party Brands, Elections, and Presidential-Congressional Relations", David R. Jones, accessed December 11, 2018, https://www.baruch.cuny.edu/wsas/academics/political_science/documents/PartyBrandsElectionsandPresidentialCongressionalRelations.pdf; Debbie Rabina and Anthony Cocciolo, *"US Government Websites During the 2013 Shutdown: Lessons from the Shutdown Library"* Alexandria: The Journal of National and International Library and Information Issues 25, 1-2 (2014): 21-30; David Scott Louk e David Gamage, *"Preventing Government Shutdowns: Designing Default Rules for Budgets"* University of Colorado Law Review 86 (2015): 181-258; Katharine G. Young, *"American Exceptionalism and Government Shutdowns: A Comparative Constitutional Reflection on the 2013 Lapse in Appropriations"* Boston University Law Review 94,3 (2014):991-1027; "Ain't No Rest for the Wicked": Population, Crime, and the 2013 Government Shutdown", Ricard Gil and Mario Macis, accessed December 13, 2019, http://repec.iza.org/dp8864.pdf; "Who Accurately Predicted the End of the Government Shutdown?", Chris C. Martin, Emory University e Kimmo Eriksson, accessed December 11, 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2609920.

[107] Regarding this point and as abovementioned, in addition the Commission, Member States may also have the right to initiate legislation in accordance with Article 76 (TFEU).

enter into force until it is approved by the Member States in accordance with their respective constitutional requirements

§ 3. Exceptions

There are cases when the Council can adopt legislation without involving the EP or the Commission at all. These situations are highly uncommon in nature and can be seen as a remnant of Council "supremacy" which was manifest in the EU's past. It is relevant to highlight Article 108(2) TFEU according to which the *"on application by a Member State, the Council may, acting unanimously, decide that aid which that State is granting or intends to grant shall be considered to be compatible with the internal market, in derogation from the provisions of Article 107 or from the regulations provided for in Article 109, if such a decision is justified by exceptional circumstances"*. On occasion, the Commission may also exercise independent legislative power in accordance with Articles 45(3), point d) and 106(3) TFEU. However, the Commission's use of this power has been rare[108].

The Commission may also be called upon to adopt Delegated or Implementing Acts (290 and 291 TFUE) [109]. Implementing Acts do "*no more than ensure the uniform implementation of the legally binding acts from which they get their legal basis. Unlike delegated acts, implementing acts cannot amend or supplement even non-essential elements of the original act. Therefore, implementing acts are well suited to regulate highly technical matters where there is a need for very specific rules to ensure harmonized implementation and application through the EU"*. Both types of acts must be expressly provided for in the original act. Delegated Acts may be revoked by Parliament or Council and may only enter into force absent objections from these Institutions. Meanwhile Implementing Acts[110] are subject to Comitology, which gives Member States the power to examine the Act (and even block it when under the examination procedure)

[108] Alina Kaczorowska, *European*…, 160; Lorna Woods and Philippa Watson, *EU Law* (11th ed., Oxford: Oxford University Press, 2012), 171.
[109] Ana Maria Guerra Martins, "Anotação ao art.º 48.º do TUE", in *Tratado de Lisboa*…, 176-182; Eileen Denza, "Article 48 [Treaty Revision Procedures]", in *The Treaty on European Union*…, 1331-1355; Luis Jimena Quesada, "The Revision Procedures of the Treaty", in *The European Union after Lisbon: Constitutional Basis, Economic Order and External Action*, Hermann-Josef Blanke e Stelio Mangiameli eds. (Heidelberg: Springer, 2012), 323-342; Miguel Gorjão-Henriques, *Direito da União: História …,* 291ff.
[110] One should note that, strictly speaking, Implementing Acts should be categorized as administrative acts and not as legislative acts.

and the European Parliament and Council the power to scrutinize it, albeit absent the blocking and revoking powers that they enjoy in Delegated Acts [111/112/113][114].

§ 4. *Passerelle* Clauses

A brief reference to the *passerelle* clauses contained within the Treaties. *Passerelle* clauses may be used to simplify or to replace the special legislative procedure originally enshrined in the Treaties with the ordinary legislative procedure. Article 48(7) of the TEU contains a general *passerelle* clause setting forth that:

a) "Where the Treaty on the Functioning of the European Union or Title V of this Treaty provides for the Council to act by unanimity in a given area or case, the European Council may adopt a decision authorising the Council to act by a qualified majority in that area or in that case. This subparagraph shall not apply to decisions with military implications or those in the area of defence" and;

b) "Where the Treaty on the Functioning of the European Union provides for legislative acts to be adopted by the Council in accordance with a special legislative procedure, the European Council may adopt a decision allowing for the adoption of such acts in accordance with the ordinary legislative procedure."

In both cases, national parliaments must be notified and can exercise veto power within six months of the notification date[115].

Additionally, there are special *passerelle* clauses applicable only to specific matters, but whose requirements are generally easier to meet than the ones for the general clause.

---

111 *See,* Michael Kaeding e Kevin M. Stack, *"Legislative Scrutiny? The Political Economy and Practice of Legislative Vetoes in the European Union"*, Journal of Common Market Studies 53, 6 (2015): 1268-1284; "A dearth of legislative vetoes: Why the Council and Parliament have been reluctant to veto Commission legislation", Michael Kaeding e Kevin M. Stack, accessed December 20, 2019, http://blogs.lse.ac.uk/europpblog/2016/10/25/a-dearth-of-legislative-vetoes/.

112 Exceptionally, in accordance with Articles 24 and 26 TFEU, the Council may adopt implementing acts (Article 291(2) TFEU),

113 *See,* "New Regulation on the rules and procedures for the operation of unmanned aircraft: Part A – Its relationship with national laws", Marília Frias and Tiago Sérgio Cabral, accessed August 20, 2019, https://www.vda.pt/pt/publicacoes/insights/new-regulation-on-the-rules-and-procedures-for-the-operation-of-unmanned-aircraft-part-a-its/21300/.

114 Under Regulation Nº 182/2011/EU of the European Parliament and of the Council of 16 February 2011 laying down the rules and general principles concerning mechanisms for control by Member States of the Commission's exercise of implementing powers.

115 *See*, António Gameiro, *O Papel dos Parlamentos Nacionais na União Europeia* (Coimbra: Coimbra Editora, 2011), 411ff.; Alessandra Silveira, "Sull'esercizio delle competenze dell'Unione europea: il Parlamento portoghese e il giudizio di conformità al principio di sussidiarietà", in *The role of national parliaments in the EU integration process* (Wolters Kluwer Italia/CEDAM: Milan, 2016); Rudolf Hrbek, "The Role of National Parliaments in the EU", in *The European Union after Lisbon*…, 129-157.

These clauses may be found in Articles 31(3) TUE, 81(3) TFUE, 153(2) TFUE, 192(2) TFUE, 312(2) TFUE and 333(1 and 2) TFUE.

There are distinct advantages in using *passerelle* clauses to replace the use of special legislative procedures by the ordinary legislative procedure. Even though one can express some concerns regarding the transparency and democratic legitimacy of the ordinary legislative procedure when taking into account trilogues, any issue pales in comparison to the ones suffered by the special procedures. The expansion of the ordinary legislative procedure under the Treaty of Lisbon was a sensible and wise decision and a simplified manner to build upon it should be welcomed. However, this is not to say that *passerelle* clauses are flawless, far from it. This legislative tool grants too much power to the ECON in the legislative procedure, where there is no apparent reason for it to have any[116]. The power to trigger a *passerelle* clause should rest with the European Parliament itself, which is the interested party and the Institution with the highest connection with the European Citizens and the highest degree of democratic legitimacy. To keep State control over the procedure, the Council could be called upon to approve or reject the triggering of the clause. Veto power by national parliaments offers a robust second layer of State control, making the systematically incoherent intervention of the ECON unneeded.

§ 5. Citizens' Initiative

Citizens can directly propose legislation at the European level, in accordance with the principles enshrined within Article 11(4) TEU and 24 TFEU. Regulation N.º 211/2011/EU of the European Parliament and of the Council of 16 February 2011 on the citizens' initiative operationalises and develops the Constitutional principles.

Using this tool entails fulfilling some strict requirements regarding representativity. The organisers must come from seven or more Member States and subscribers from, at

[116] The reason is probably due to the fact that using a *passerelle* clause effectively changes the basic rules of European Constitution law as enshrined in the Treaties and, in accordance to the underlying logic, this should be left to heads of State in the European Council. But that logic is incredibly flawed. First, opinions in the Council and European Council will be the mirror image of each other more frequently than not, so needing the European Council to trigger the clause and the Council to approve the (future) legislation is not logical. If we have to choose between the two institutions, the Council, which actually was law-making powers would be the most adequate. Second, those heads of State will frequently not have any law-making, let alone constitutionally amending powers on their States, it does not seem coherent for them to have it at a European level. Furthermore, they lack the mandate or democratic legitimacy for it. The fact the European Council is not needed is even clearer when national parliaments (which tend to have a higher degree of legitimacy) are involved and have veto power, therefore being perfectly able to raise any issues that a Member State might have with the amendment.

least, ¼ of the Member States (Article 3(2)). The minimum number of signatures is 1 million with distribution rules. Besides the ¼ of the Member States rules, organisers have to ensure that signatures, also, correspond, at least "*to the number of the Members of the European Parliament elected in each Member State, multiplied by 750*" (Article 7(2)).

The Commission has the power to stop initiatives at an early stage by refusing their registration. However, this power should only be used when: *i)* the initiative manifestly falls outside the framework of the Commission's powers to submit a proposal for a legal Act of the Union for the purpose of implementing the Treaties; *ii)* the initiative is manifestly abusive, frivolous or vexatious; or *iii)* the proposed citizens' initiative is manifestly contrary to the values of the Union as set out in Article 2 TEU[117].

Displaying a high degree of attention to the issue of democratic participation, the legislator opted to guarantee that *"technology [is put] to good use as a tool of participatory democracy"* by allowing signatures to be collected interchangeably in paper form or electronically.

If success is achieved by the organisers and the initiative meets its goal, the Commission must publish it in the register with celerity *and "receive the organisers at an appropriate level to allow them to explain in detail the matters raised".* Furthermore, the Commission, within three months, has to *"set out in a communication its legal and political conclusions on the citizens' initiative, the action it intends to take, if any, and its reasons for taking or not taking that action"* (Article 10).

Additionally, a public audience must be organised, also within three months, where organisers have the opportunity to explain and argue for their proposal. The audience shall be conducted in the EC and, in addition to the Parliament and Commission, is open to the participation of other institutions and bodies that show interest[118].

[117] The General Court was called upon to intervene in two occasions when the Commission refused citizen's initiatives in an unlawful manner. *See*, Pedro Infante Mota, "Acordos Mega-Regionais", *in União Europeia – Reforma ou Declínio*, Eduardo Paz Ferreira coord. (Lisboa:Vega, 2016), 376-399; Judgment of EGC of 3 February 2017, *Minority SafePack - one million signatures for diversity in Europe*, Case T-646/13, ECLI:EU:T:2017:59; Judgment of the EGC of 10 May 2017, *Stop TTIP (Efler v. Commission)*, Case T-754/14, ECLI:EU:T:2017:323; Judgment of the EGC of 24 September 2019, *Romania v. Commission*, Case T-391/17, ECLI:EU:T:2019:672.

[118] Patrícia Calvão Teles, "Anotação ao artigo 24.º do TFUE", in *Tratado de Lisboa Anotado e Comentado,* coords. Manuel Lopes Porto and Gonçalo Anastácio (Coimbra: Almedina, 2012), 264-267; Dulce Lopes and Paula Veiga, "Anotação ao artigo 11.º do TUE", in *Tratado de Lisboa Anotado e Comentado,* coords. Manuel Lopes Porto and Gonçalo Anastácio (Coimbra: Almedina, 2012), 54-57

Chapter III – Current Trends for AI in the European Union

§ 1. Preliminary Work

§ 1.1. European Council Conclusions – 19 October 2017

The European Council Conclusions of 19 October 2017 (hereinafter, "ECON's Conclusions"), albeit short and, due to their nature, providing much less detail when compared to some of the documents we shall examine below, deserve to be mentioned, due to the political influence of this Institution, responsible for defining the Union's overall political direction and priorities.

In Part II of the ECON's Conclusions, dedicated to a Digital Europe, the European Council considers that it is paramount to: *a)* modernize governments and the public sector; *b)* speed up the implementation of legislative initiatives in within the Digital Single Market[119], *c)* [invest in the] improvement of the infrastructure and communications network, which includes the roll-out 5G across the EU and the freeing up of spectrum frequencies, *d)* [implement] a common and comprehensive approach to cybersecurity; *e)* intensify the fight against online crime and terrorism; *f)* [adapt] the education system and prepare the labour market for the changes brought by technology and; *g)* urgently address emerging trends like artificial intelligence and blockchain technology. On the last point the European Council calls on the Commission to *"to put forward a European approach to artificial intelligence by early 2018 and to put forward the necessary initiatives for strengthening the framework conditions with a view to enable the EU to explore new markets through risk-based radical innovations and to reaffirm the leading role of its industry"*.

[119] We would like to point out that at the time of writing all the initiatives that were given priority by the European Council under the "future-oriented regulatory framework" have either been implemented or have been agreed and are under in the process of being implemented. Those include the Geo-Blocking Regulation, the new Audio-Visual Media Services Directive, the Copyright Directive, the New Electronic Communications Code and the Regulation on the free-flow of non-personal data.

§ 1.2. The European Parliament's Legal Affairs Committee European Civil Law Rules in Robotics Study[120]

While a profound analysis of the conclusions of this study is hardly warranted, as it is outdated and lost relevance after the Parliament's Recommendation that would proceed it, reading the Civil Law Rules in Robotics Study (hereinafter, "Robotics Study") one can easily see its influence on the subsequent Parliamentary actions on this issue.

The Robotics Study was commissioned on the basis of the *Draft Report* that served as the basis for the Recommendation. The Robotics Study had the merit of pointing out some flaws in matters like the definition of autonomous robots which were largely ignored by Parliament[121]. However, the study did not address flaws related to the overlooking of AI itself in favour of robotics (*see* point below). True the definition is coherent, but also coherently badly applied.

Instead of that, the Robotics Study takes issue with the concept of smart robot due to an idea it christens as western fear of the robot[122]. The suggestions given for debunking baseless fears and keeping AI under human control are relevant for sure and would go on to be mirrored in other instruments by European Institutions. However, one cannot stop oneself from thinking that the study itself suffers from a relevant degree of bias, even if it is not intentional. *Astro Boy* and others like *Doraemon* and *Mega Man* are, indeed, quite influential in Japanese culture. But if we follow that path to the obvious conclusion, we can argue that Japanese people do prefer to have the robots well under their control like they do in *Neon Genesis Evangelion, Mobile Suit Gundam, Tengen Toppa Gurren Lagann, Code Geass* and others.

Moreover, it is wrong to tell that "good" robots do not exist in western culture. A short list of examples would be R2-D2 (Star Wars), Marvin The Paranoid Android (The Hitchhiker's Guide to the Galaxy), Data (Star Trek), K-9 (Doctor Who), The Final Five (Battlestar Galactica), Transformers (who were created in the East but are, at least, as

---

[120] *See*, "European Civil Law Rules in Robotics" Nathalie Nevejans (requested by the European Parliament's Committee on Legal Affairs), access January 10, 2019, http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU(2016)571379_EN.pdf.

[121] The Study's critique of the use of the "independently of external control or influence" seems quite pertinent. The word influence is glaringly too imprecise to use in this context. Even the more advanced AI's will be influenced by external data, namely by the data they receive from the environment, as humans are.

[122] On why we may be afraid of robots *see*, Michael Szollosy, "Why Are We Afraid of Robots? The Role of Projection in the Popular Conception of Robots", *in Beyond Artificial Intelligence: The Disappearing Human-Machine Divide*, Jan Romportl, Eva Zackova and Jozef Kelemen eds. (Cham: Springer, 2015), 121-131.

popular in the West), the Iron Gian and Wall-E[123]. Even if the concept is quite debatable, it does not appear to be having an effect on the speed in which AI is developed across nations. The US should have inherited the Western Fear of the Robot, but it is still the number one potency in AI development as we will see in our chapter on the United States.

The Robotics Study's suggestions regarding the impact of the AI in the labour market are generally balanced, especially in what concerns the list of competences for the proposed European Agency. Giving the European Agency the ability to prepare socio-economic assessments appears to be a good suggestion and, even if not expressly stated, can be subsumed to the final list of competences proposed for the Agency.

The Robotics Study also makes pertinent proposals regarding the way in which Asimov's Laws are included in the *draft report*. Fact is that the redaction did not change much in the Resolution itself and it is clear the Asimov's law cannot be the basis of a legal or even self-regulatory framework even if you can extract some important principles from them (albeit the part that deals with ethics was improved somewhat diminishing the problem).

The Study is genuinely against the idea of giving any kind of legal personality to AI. We tend to agree, as it will be possible to see in Part II, specifically on our chapter on Product Liability. Our agreement is time-restricted though, for now, and in the context of Narrow AI (more in our chapter on the Parliament's Resolution and Part II, on Product Liability). We certainly do not share the apocalyptic vision that giving a legal personality to AI akin to what is given to companies would be equal to *"tearing down the boundaries between man and machine, blurring the lines between the living and the inert, the human and the inhuman. Moreover, creating a new type of person – an electronic person – sends a strong signal which could not only reignite the fear of artificial beings but also call into question Europe's humanist foundations. Assigning person status to a nonliving, non-conscious entity would therefore be an error since, in the end, humankind would likely be demoted to the rank of a machine. Robots should serve humanity and should have no other role, except in the realms of science-fiction"*[124].

---

[123] Again, this does not mean that we do not believe that the utmost care is needed when dealing with AI, as we had to opportunity to express before. However, the cultural arguments used in the study do not seem to hold under scrutiny. *See*, Tiago Sérgio Cabral, *"Robotics and AI in the European Union: opportunities and challenges"*, UNIO - EU Law Journal. 4, 2 (2018): 135-146.

[124] We chose to directly quote this excerpt because it reflects the unintentional bias of the study in fact, goes dangerously close to fear mongering. Fact is that in the US companies (which have legal personality based on a legal fiction) were given certain fundamental rights that in the European optic should be exclusive to natural persons. First, companies have the right to free speech, and according to Supreme Court of the United States in Citizens United v. the Federal Election Commission that entails the right to fund advertising either in favour or against a candidate. Furthermore, in Burwell v. Hobby Lobby Stores, Inc the Supreme

Back to the problem arising in the world liability, the Robotics Study does point out some interesting points that are still currently not solved. One of those is the issue of open-source software and how can we protect consumers while also fostering a welcoming environment for open-source development (which we will see below is one of the priorities of the European Commission). The second issue and directly related to ML is the issue of liability for information that was learned by the AI. While the study points that the situation addressed is *"robot causes any damage when in use or while still learning"* we must clarify that certain types of AI-powered robots will, most likely, keep learning permanently while in use.

The Study goes on to suggest the integration of the following principles of ethical principles for AI[125]:

a) Protecting humans from harm caused by robots;

b) Respecting the refusal of care by a robot;

c) Protecting human liberty in the face of robots;

d) Protecting humanity against privacy breaches committed by a robot;

e) Managing personal data processed by robots;

f) Protecting humanity against the risk of manipulation by robots;

g) Avoiding the dissolution of social ties;

Court of the United States decided companies do not have to provide their female employees with free contraceptives if doing so goes against its (their owner's) religious freedom. In the second case, with all its nuances, the court seems to have accepted that this religious freedom should prevail when in conflict with the employee's right to human dignity, to health and her religious freedom. With this in mind, and even if we disagree with this line of thinking, we are of the opinion that the line between what is human and what is a legal fiction was not weakened, neither did we (or any person from the United States that we know of) gain sudden doubts about what is living and what is inert and what is human and what is inhuman. Furthermore, we would like to remember that this study was prepared taking into account the paradigm of Narrow AI, not a future General AI. It is borderline condescending to state the European citizens will not be able to differentiate between what is human and their Google Now, Tesla or Roomba just because it has a legal personality based on a fiction. If it were a General AI, with capabilities very close to a human, then we would have a different set of legal and ethical questions and even the question of robots being only for serving humanity would be doubtful as it could be seen as enslaving intelligent life. Of course, this does not mean that we agree with suggestion of giving legal personality to AI (we do not), and we shall study it in more detail. However, if the solution is to be rejected it should not be on weak and borderline fallacious grounds. It should be because there are better solutions and, in fact, before losing itself in this kind of undesirable argumentation the Study does provide a proper reason such as the fact that maybe it is too complicated a solution for a result that might achieved in a less complicated manner. *See*, Judgment of the Supreme Court of the United States of 21 January 2010, *Citizens United v. the Federal Election Commission*; Judgment of the Supreme Court of the United States of 30 June 2014, *Burwell v. Hobby Lobby Stores, Inc*; Tiago Sérgio Cabral, *"Testemunhas de Jeová e a Liberdade Religiosa no séc. XXI: Uma Análise com base no Acórdão Palau-Martinez vs. France"*, e-Pública: Revista Eletrónica de Direito Público 4,2 (2017): 196-219.

[125] As we stated above the Study does suffer from the same issue as the Resolution (below) in that it frequently talks about robotics when it should be really talking about AI.

h) Equal access to progress in robotics and;

i) Restricting human access to enhancement technologies.

These principles are quite well reflected on the Recommendation and thus we shall address them there.

Lastly, the Robotics Study focuses on the question of the proposed Charter of Robotics (see below).

§ 1.3. The European Parliament's Civil Law on Robotics Resolution[126]

As previously stated, when addressing legislative procedure in the EU, even though the European Parliament has no power to propose legislation by itself, this Institution can request the Commission's intervention when it deems necessary. This is what it did with the Civil Law on Robotics Resolution (hereinafter, "The Resolution").

In its initial considerations, the European Parliament quotes Mary Shelley's Frankenstein's Monster, the classical myth of Pygmalion, the story of Prague's Golem and the robot of Karel Čapek to argue that humanity's wish to build intelligent machines that can serve as companions to humans is an old desire[127].

The Resolution considers that we are finally on the verge of achieving this long-held dream. However, according to the European Parliament, AI is not completely rose-coloured and will bring with it plenty of challenges. The Resolution highlights the challenges to the labour market, the judicial system, data protection and ethics. There is even a nod to the (conceivable) scenario where AI can eventually surpass human intelligence.

Parliament starts by arguing that a *"series of rules, governing in particular liability, transparency and accountability, are useful, reflecting the intrinsically European and universal humanistic values that characterise Europe's contribution to society, are necessary; whereas those rules must not affect the process of research, innovation and development in robotics"*. One should note that this desire to have an AI with humanistic values without hindering its development is extremely difficult

[126] European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics, http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0051+0+DOC+XML+V0//EN

[127] We gave a few more examples on our introductory chapters on AI.

to fulfil. But the European Parliament did not choose to prioritise values over the economy or vice-versa, what it asked the Commission was to find a solution and manage to combinate both. Furthermore, Parliament believes that the EU should play a key role in creating these ethical standards and they should be reflected in its Regulations and Codes of Conduct, arguing for a solution anchored in more European cooperation.

The Resolution covers a large number of issues, and while not offering in-depth coverage of any specifics, it does offer valuable insight into what the current sensibilities are in this Institution. The matters addressed are: a) General principles concerning the development of robotics and artificial intelligence for civil use; b) Research and innovation; c) Ethical principles; d) Creation of a European Agency; e) Intellectual property rights and the flow of data; f) Standardisation, safety and security; g) Autonomous means of transport; h) Care robots; i) Medical robots; j) Human repair and enhancement; k) Education and employment; l) Liability and; m) International aspects. Some of these have their own chapters. Where that is the case, we will confine ourselves to describe and shortly comment the Parliament's position.

§ 1.3.1. General Principles Concerning the Development of Robotics and Artificial Intelligence for Civil Use

Any AI regulation at the European level must start by defining what we wish to regulate. That is why common definitions are a must. Albeit, the European Parliament clearly understands this, the fact is that, out of the gate, there are deep flaws in their proposal for common Union definitions of cyber-physical systems, autonomous systems and, smart autonomous robots. Truth be told this is a flaw stems not necessarily from lack of technical expertise but from the fact that the Resolution, in general, is too focused on robots and overlooks AI in general. A robot does not have to be AI-equipped and AI does not need to be externalised through a robot. The fact that AI is not supported through a robot does not make it less in need of a framework regulating it[128].

Let us see some examples: a) A phone's personal assistant; b) An algorithm that is able to detect diseases when you input the results of the patient's tests; c) An algorithm

[128] Taking as an example the abovementioned definition of AI by the HLG and the adaptations introduced within it by us, AI as scientific discipline includes "*includes several approaches and techniques, such as machine learning (of which deep learning and reinforcement learning are specific examples), machine reasoning (which includes planning, scheduling, knowledge representation and reasoning, search, and optimization), and robotics (which includes control, perception, sensors and actuators, as well as the integration of all other techniques into cyber-physical systems".*

that decides if the client of a bank should be given a loan based on the automatic combing of all available financial records; d) An algorithm that through the use of CCTV images is able to detect potential criminals in real-time; e) An algorithm that is able to predict recidivism of a criminal based on his/hers previous history and data from his/her personal life; f) An algorithm that controls the nuclear arsenal of a country and is able to comb the internet in search of threats and automatically fire if it finds an immediate threat that cannot be addressed by conventional means and; f) a Roomba vacuum cleaner. According to the criteria proposed by the EP, the only one of these examples that is a smart robot and needs regulation by the EU is the Roomba vacuum cleaner. It fits every one of proposed criterion namely: *i)* the acquisition of autonomy through sensors and/or by exchanging data with its environment (inter-connectivity) and the trading and analysing of those data; *ii)* self-learning from experience and by interaction (optional criterion); *iii)* at least a minor physical support; *iv)* the adaptation of its behaviour and actions to the environment and; *v)* absence of life in the biological sense. Most other examples fail in adapting their behaviour and actions to the (physical) environment and potentially – depending on what you consider to be the acquisition of autonomy – on the first criterion.

Of course, this is if you consider that when the Resolution talks about the environment it is the physical environment. However, if is not and it includes, for example, the digital environment, the flaws run more profound, because than everything is an Autonomous Robot. Is your personal computer a robot in this scenario? Well it does depend on what you consider to be autonomous (and the Resolution does not help there) but let us go for a broad interpretation of the concept as a way of showing our reluctance regarding the criterion of the EP. *i)* Your computer possesses a plethora of sensors like your microphone and webcam, and it can exchange data with environment through the internet (or through its sensors) and trade and analyse that data; *ii)* It might not have the capability of self-learning from experience and by interaction but that is an optional criterion, or it might have if you have Microsoft's Cortana, Apple's Siri or Google Now; *iii)* it has, at least, minimal physical support; *iv)* it can adapt its behaviour and actions to the environment. It can, for example, (hopefully) stop viruses and stop downloading updates if it detects a slow internet connection and; *v)* it is not alive in a biological sense. Seems that you already have an autonomous Robot in your hands. Maybe it is time to get some new Regulation then, but we do not believe that this was Parliament's intention.

In conclusion, the proposed definition overlooks AI in favour of Robotics, it is not precise enough, being either too broad or too strict. Furthermore, in the context of the Resolution some concepts such as autonomy should be defined and are not.

Still on the issue of general principles the Resolution's focus, as expected, the Single Market. This no more than natural. If we look into the exclusive and shared legislative competences of the EU (Articles 3 and 4 TFEU), it is expectable that a significant part of the Union's intervention is based on the Single Market by itself or jointly with another competence such as environment, transport or consumer protection. The exception could be if it were necessary to establish competition laws specific for AI, but even then this exclusive competence would probably appear jointly with the single market (the competence to establish said rules is granted to the EU so it can safeguard the market after all).

The Resolution requests uniform rules for introducing autonomous robots in the single market. It also calls for registration of certain types of autonomous robots (though it does not specify which types) and stresses importance of start-ups and small and medium-sized enterprises.

Last but not least, it calls for human control over autonomous robots and AI, stressing that humans should not be replaced by AI. It also stresses that particular caution should be exercised regarding AI-Human relationships[129].

[129] Relationships and love between humans and AI are not exactly new phenomena in our collective imagination. Humans are creatures with a flair for the dramatic though and, as such, we like to predict how said relationships will fail. In the 2013 movie "Her" a man falls in love with an AI (operative system) but the relationship eventually fails due to AI's superior intelligence and inclination to polygamy. In the Japanese anime "Plastic Memories" the reciprocal love between young Tsukasa and the android Isla ends tragically when she starts to degrade at the end of her lifecycle. Though some fascination appears to exist in finding love and living happily ever after with our artificial counterparts, current applications of the technology appear to focus on the more lustful aspects of the relationship. We have developed such "interesting" appliances as AI-equipped sex dolls and AI-equipped masturbatory aids. In addition to the ethical implications that may arise from these uses of AI, we would like to point out the data protection concerns that we have regarding them, as frequently data is collected from users and used to "perfect" the machine. *See*, "This company specialises in talking, AI-powered sex dolls", BBC News, accessed July 10, 2019, https://www.bbc.com/reel/video/p06f6xn2/this-company-specialises-in-talking-ai-powered-sex-dolls; "Don't Get Your Valentine an Internet-Connected Sex Toy", Emily Dreyfuss, accessed July 10, 2019, "Autoblow AI is a sex toy that promises 'surprise'", Daniel Cooper, accessed September 30, 2019, https://www.engadget.com/2019/09/27/autoblow-ai-deep-learning-sex-toy/.

§ 1.3.2. Research and Innovation

In this area, the European Parliament asks for more cooperation between Member States and more funding by both Member States and the European Union. The proposal for the new multiannual financial framework already started addressing this question[130] along with the Communications from the EC and deliverables from the HLG which we will see in the sections below. Parliament also addresses the need for proper infrastructure and to keep the Union's commitment to net neutrality, to open science and responsible ethical innovation and to interoperability, open standards, open development environments, open platforms, transparency and innovative (and fair) licensing models.

§ 1.3.3. Ethical Principles

Ethics in AI will be a recurrent theme in the European Institutions' positions about the issue. It is not surprising to see that it is one of the highlights of the Resolution.

The applicability of certain fundamental principles of European Union law to AI is restated, such as human dignity, equality, justice and equity, non-discrimination, informed consent, private and family life and data protection, non-stigmatisation, transparency, autonomy, individual responsibility and social responsibility, and on existing ethical practices and codes. Specific principles for AI are also called upon, such as beneficence, non-maleficence, autonomy and justice. Further details about the interpretation given to these principles by Parliament are contained in the Charter of Robotics that serves as an annex to the Resolution (which we find suffers from not focusing on AI and instead on robotics and not being more detailed into how its principles can be programmed into AI, if the necessity arises).

One must note how prominent the principles regarding data protection and privacy are in this context. In addition to being referred along with the general principles, two special paragraphs are dedicated to Robots that are placed in *"traditionally protected and private spheres because they are able to extract and send personal and sensitive data"* and to transparency. Transparency in this sense can be reduced to "GDPR transparency" of which we shall talk in further detail further (*see* Part II, our chapter about the GDPR), but for now an acceptable simple explanation is that humans should always be able to understand

[130] *See,* section 2.1. "Communication from the European Commission: Artificial Intelligence for Europe"

decisions taken by autonomous means, including AI if they have a relevant impact on their life. Parliament considers that a "black box" should be installed in AI-equipped machines to track and, if needed, restore their decision process. The choice of words is quite unfortunate since the aviation-style black box shares its name with the AI style black box (a notion we have explained in our introductory chapters).

Lastly, Parliament calls for an updated legal framework guided by these legal and ethical principles and that takes into account the complexities surrounding AI.

§ 1.3.4. Creation of a European Agency

The European Parliament requests the creation of a European Agency for Robotics and Artificial Intelligence. Parliament believes that this agency should have an advisory role to both Member States and the European Union "*identifying standards for best practice, and, where appropriate, recommending regulatory measures, defining new principles and addressing potential consumer protection issues and systematic challenges*" and reporting on the relevant developments.

While the creation of Agency seems like a pertinent measure, its description (again) suffers from highlighting excessively the robotics part of this intended body and overlooking AI. It starts with the name the agency itself. Ideally, it should be called European Agency for Artificial Intelligence or European Agency for Artificial Intelligence and Robotics. AI is the broader field of study and reason behind the creation of the agency. If not for AI and, with it, AI-equipped robots this agency would not have a *raison d'être*. One might even question oneself if there is a reason for this Agency to occupy itself with non-AI equipped robots. Contrarily so, we are reasonably sure that there are very good reasons for the proposed Agency to address questions related to non-robotic AI.

This should not be seen as petty criticism regarding the Agency's name. In accordance with what we have explained before it shows a deeper misunderstanding in the Resolution, and that can have unpredictable consequences. For example, according to the European Parliament the Agency must help in *"ensuring a timely, ethical and well-informed response to the new opportunities and challenges, in particular those of a cross-border nature, arising from technological developments in robotics, such as in the transport sector"*. Parliament was clearly (too) focused on autonomous vehicles, but AI has other, arguably more life-changing uses, that are not necessarily dependent on Robots (selling and buying stocks and potentially crashing

the market if things go wrong or managing a country's entire war arsenal, for example). Should we save our timely, ethical and well-informed responses to situations where there is a Robot and not where there is, for example, a piece of software running in the cloud? We do not think so, and probably neither does the European Parliament. Thereby, any forthcoming initiative or legislative proposal must be careful to avoid the same mistake. On this note, the reporting duties of the Agency are described as only concerning robotics, which we think is a mistake that originates from the same source and should be addressed in the same manner.

Furthermore, in the context of specific AI legislation we favour and agency with regulatory powers, supported by national regulators instead of just an advisory role. Said agency should also work with stakeholders to foster the development of AI in the EU through initiatives such as regulatory sandboxes.

§ 1.3.5. Intellectual Property Rights and the Flow of Data

Data Protection and Privacy are (again) the highlights of this specific point. In an abbreviate manner, (because we will study this a much more profound manner in our chapter regarding the GDPR), the European Parliament considers that the development of AI should respect both the principles of the GDPR and the general fundamental of Privacy, Data Protection and Freedom to Conduct Business.

Some priorities of the Digital Single Market are reasserted, such as data protection and cybersecurity. These priorities are the reason for digital single market legislation like the GDPR, the Regulation 2018/1807/EU 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 (Regulation on the free flow of non-personal data), the Directive 2016/1148/EU of the European Parliament and of the Council of 6 July 2016 concerning measures for a high common level of security of network and information systems across the Union (NIS Directive), the Cibersecurity Act[131] and the Open Data and Public Sector Information Directive[132].

[131] Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013.
[132] Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information

On intellectual property itself, the EP only states *that "there are no legal provisions that specifically apply to robotics, but that existing legal regimes and doctrines can be readily applied to robotics, although some aspects appear to call for specific consideration"* and that the EC should support a technological neutral approach.

§ 1.3.6. Standardisation, Safety and Security

It is quite interesting how safety and the single market are intertwined in the European Institutions' general AI strategy. Here we see the fostering of international and open standards with two main objectives: *i)* ensure that AI is safe for the consumer; *ii)* ensure healthy competition in the field of AI by preventing locking-out.

The issue of testing is also, albeit lightly, addressed. Parliament proposes little more than a harmonised European framework for AI-related testing.

§ 1.3.7. Autonomous Means of Transport

§ 1.3.7.1. Autonomous Vehicles

Parliament starts by recalling that the concept of autonomous means of transport is not restricted to autonomous cars. It also includes air transport, rail and waterborne transport. Though, it is in the automotive industry that this Institution believes the most urgent needs for new and improved regulation exist since a fragmented approach could harm the single market and hinder the implementation of this technology in the European Union.

The possible need for driver takeover of an automated vehicle is also examined by the European Parliament. However, one cannot help but get the feeling that the European Parliament's Resolution does not take into account the different levels of automation that an automated vehicle can possess. In fact, it can go from simple driver assistance to complete automation where there is no need for the driver to take control, ever. The Society of Automotive Engineers "J3016" standard, used by the U.S. Department of

Transportation contains five levels of automation, with the last two being considered full automation[133].

Challenges brought by full automation differ from challenges brought by partial automation. For example, overreliance on semi-autonomous systems must be avoided because they are not designed to completely take over the driver. We do believe that this was the issue that most bothered Parliament when writing this Resolution, and it is not without reason. There have been some reported cases of people giving complete control to Tesla's semi-autonomous driving system. Since the system is not designed to replace the driver, this can give rise to potentially dangerous and mortal situations (no reported accidents as far as we know, though). What Parliament should have done was to call on manufacturers of semi-autonomous systems to study not only how much time the driver takes to react but also means to avoid over-reliance on the system. To their credit Tesla's system only works when the driver has, at least, one hand on the wheel, but that did not prevent some drivers from, in a fantastic display of human ingenuity, falling asleep with said hand properly placed or using some other artifices to fool the technology[134].

Things are different for a fully autonomous system, where the autonomous vehicle shall perform at least at the same level as a human driver, and probably better. In these conditions there is no need for the driver to be prepared to take over the wheel because doing so will probably result in worst overall performance. This does not mean that other challenges do not exist, and Parliament pointed some like liability, road safety, environmental issues, data and ICT infrastructure-related challenges and the necessary adaptation to the labour market. But the consequences are different for semi-autonomous or autonomous system. For example, regarding liability, if a system depends on the user

---

[133] *See*, SAE International, *J3016 – Jun 2018 Standard - Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles* (Pennsylvania/Michigan: SAE International, 2018); US Department of Transportation, *Preparing for the Future of Transportation: Automated Vehicle 3.0* (Washington, DC: US Department of Transportation, 2018, p. 6ff.; "Automatic Driving Systems: A Vision for Safety", US Department of Transportation, access April 5, 2018, https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/13069a-ads2.0_090617_v9a_tag.pdf.

[134] *See*, "Video appears to show Tesla driver asleep at the wheel", The Guardian Staff, accessed September 11, 2019, https://www.theguardian.com/technology/2019/sep/10/video-appears-to-show-tesla-driver-asleep-at-the-wheel-car; "A Sleeping Tesla Driver Highlights Autopilot's Biggest Flaw", Alex Davies, accessed September 11, 2019, https://www.wired.com/story/tesla-sleeping-driver-dui-arrest-autopilot/; "A new video shows a Tesla driver who appears to be asleep while driving down the Massachusetts Turnpike", Lisa Eadicicco, accessed September 11, 2019, https://www.businessinsider.com/tesla-driver-asleep-while-driving-down-massachusetts-turnpike-video-2019-9; "Tesla drivers are getting caught sleeping on Autopilot – blame people, not Autopilot", Fred Lambert, accessed September 11, 2019, https://electrek.co/2019/06/16/tesla-driver-caught-sleepingn-autopilot-blame/; "Fully Sleeping" Tesla Driver Cruises 30 Miles on Autopilot", Kristin Houser, accessed September 11, 2019, https://futurism.com/the-byte/sleeping-tesla-driver-autopilot.

taking control and the user for some reason does not we have to see if this was a problem of the user not reacting on time or of the system not giving the user enough time to react (Parliament somewhat addressed this). However, if the system is fully autonomous that is not an issue, and the user will, probably not be liable. Therefore, the problem here is discovering what failed on the system. Regarding the labour market, Parliament points the need to train heavy goods vehicles drivers for the use of autonomous vehicles, and indeed the abovementioned need exists, but only for semi-autonomous vehicles. For completely autonomous vehicles, the problem is more in line with the general changes in the labour market brought by AI.

Fortunately, we can get a sneak peek regarding the introduction of automation and its advantages by looking at another industry: aviation. Currently, vital functions of airplanes are performed by a computer with a high degree of effectiveness (European manufacturer Airbus is particularly supportive of automation). In fact, airplanes are the safest mode of transportation and most accidents are caused by human error. Of course, planes are not entirely automated, but they are getting there, and since driving a car is less complicated than challenging nature itself by soaring through the air while piloting an airplane[135], autonomous vehicles will be probably be mainstream technology before fully autonomous airplanes.

Lastly, Parliament stresses the need to complete EGNOS and Galileo as to ensure that the European Union is not dependent on foreign satellite networks and that providers have a competitive European solution they can choose for the provision of their services. The anticipated benefits of this technology for people with reduced mobility are also brought to attention.

§ 1.3.7.2. Drones

Drones or unmanned aerial vehicles do not need to be AI-equipped to represent a regulatory challenge[136]. However, Parliament's call on the Commission to provide further

[135] Also, the average commercial pilot is a highly trained professional, so it is much more challenging for a machine to surpass him/her than it is to surpass the average driver. Machines have been winning against average players in videogames for quite some time, but beating champions is a more recent achievement (sometimes with decades separating the two).

[136] *See*, " A passenger jet pilot swerved to avoid drone near Gatwick Airport", Rob Picheta, accessed August 30, 2019, https://edition.cnn.com/2019/08/28/uk/gatwick-drone-near-miss-scli-gbr-intl/index.html; "Airports scramble to handle drone incidents", Matt McFarland, accessed August 30, 2019, https://edition.cnn.com/2019/03/05/tech/airports-drones/index.html; "Gatwick Airport: Drones ground

regulatory densification on the issue appears to out of context in the Resolution. It seems like to reference was introduced in the Resolution more to remember the Commission to take action and exercise the powers given to it by Regulation 216/2008/EC[137]. In the meantime Regulation 216/2008/EC was repealed and replaced by Regulation 2018/1139/EU of the European Parliament and of the Council of 4 July 2018 on common rules in the field of civil aviation and establishing a European Union Aviation Safety Agency, and amending Regulations (EC) No 2111/2005, (EC) No 1008/2008, (EU) No 996/2010, (EU) No 376/2014 and Directives 2014/30/EU and 2014/53/EU of the European Parliament and of the Council, and repealing Regulations (EC) No 552/2004 and (EC) No 216/2008 of the European Parliament and of the Council and Council Regulation (EEC) No 3922/91 (hereinafter, "Basic Regulation"). The Commission did exercise its powers (as attributed by the new Basic Regulation) through Commission Implementing Regulation 2019/947/EU of 24 May 2019 on the rules and procedures for the operation of unmanned aircraft and Commission Delegated Regulation 2019/945/EU of 12 March 2019 on unmanned aircraft systems and on third-country operators of unmanned aircraft systems[138].

§ 1.3.8. Care Robots

While conceding that care robots are getting more advanced and less expensive to produce and that currently they have a wide range of application, potentially helping the elderly, people with disabilities or suffering from dementia, cognitive disorders, or memory loss, the European Parliament draws attention to the fact that they cannot replace care given by humans and/or provide the essential social interaction and that trying to do so would dehumanize care.

flights", BBC News, accessed August 30, 2019, https://www.bbc.com/news/uk-england-sussex-46623754; "Gatwick returns to normality but drone threat remains", Jamie Grierson, accessed August 30, 2019, https://www.theguardian.com/world/2019/jan/04/gatwick-returns-to-normality-but-drone-threat-remains.

[137] Regulation 216/2008/EC of the European Parliament and of the Council of 20 February 2008 on common rules in the field of civil aviation and establishing a European Aviation Safety Agency, and repealing Council Directive 91/670/EEC, Regulation (EC) No 1592/2002 and Directive 2004/36/EC.

[138] *See*, "New Regulation on the rules and procedures for the operation of unmanned aircraft: Part A – Its relationship with national laws", Marília Frias and Tiago Sérgio Cabral, accessed August 20, 2019, https://www.vda.pt/pt/publicacoes/insights/new-regulation-on-the-rules-and-procedures-for-the-operation-of-unmanned-aircraft-part-a-its/21300/.

§ 1.3.9. Medical Robots

Parliament focuses on three aspects regarding medical robots: *i)* ensuring that professionals are trained to work with them; *ii)* keeping the final decision with humans and, *iii)* making available the adequate testing procedures to guarantee the safety of medical robots.

Regarding the first aspect, the EP considers that if proper training is not given to the professional that will use the medical robots, the health of the patients can be harmed. Minimum professional qualification requirements should, therefore, be imposed for professionals before allowing them to use this type of devices. With the growing trend of self-diagnosis – "Doctor Google" and its propensity for diagnosing fatal diseases when you have a simple flu is a notorious example – Doctors must be prepared to deal with these types of situations.

Medical robots should help Doctors and not replace Doctors in the view of the European Parliament. Thus, the final say about the medical treatment administered to a patient should always lie with the physician. The Doctor-Patient relationship should not be weakened by the appearance of the robotic player.

Parliament also urges action by the Commission to ensure safe testing conditions for medical robots, *"particularly in the case of devices that are implanted in the human body, before the date on which Regulation 2017/745/EU on medical devices becomes applicable"*.

With all this in mind, Parliament believes that, if rightly used, medical robots can raise life expectancy, enhance quality of life, improve outcomes in rehabilitation, provide highly effective logistical support within hospitals and reduce the cost of healthcare.

§ 1.3.10. Human Repair and Enhancement

Pragmatic. That would be the best way to describe the European Parliament's position on human repair and enhancement. Parliament does not appear to harbour preconceptions regarding integration of mechanical parts on a human body for repairs or even for plain enhancement. That notwithstanding, it still draws attention to the fact that new and highly complicated ethical issues will arise and there is the need to prepare for them. Furthermore, it warns that there is the need to ensure that maintenance is done, and updates are made available for persons equipped with cyber-physical systems. Parliament

proposes that independent entities should be created to guarantee basic upkeep even if the original supplier/manufacturer is unable or unwilling to do so. Manufacturers would have to supply comprehensive design instructions including source code to these independent entities. Cybersecurity is also a key priority that must be addressed. Lastly, guaranteeing equal access to these treatments and enhancements should be a priority for the European Union.

§ 1.3.11. Education and Employment

There is nothing similar to a scientific consensus on this issue and Parliament, smartly one might add, focuses on current problems and calls on the Commission to be on the lookout for future developments.

On the short-term issues, we have lack of ICT skills in the European workforce and the need to both provide adequate teaching for current students that will be part of the workforce in the near future and to train those who are already part of the workforce but risk being made "obsolete" by emerging technologies.

Lack of women in ICT related fields of study has been a problem ever since said ICT related fields were created[139]. It must be stressed that Parliament could have given more prominence to this particular matter, in fact, it could even have been detached from Education and Employment[140]. It is indeed a concern for the labour market but more than that, it is part of broader equality problem – more accurately, the deficit of thereof. Equality between men and women is a fundamental right enshrined in the Treaties, the Charter of Fundamental Rights of the European Union and in the shared constitutional traditions of Member States and as such merits proper consideration.

Long term, Parliament calls on the Commission to beware of changes in the job market and to act with the objective of ensuring that the changes brought by AI are harnessed in a way that is positive for the general workforce and does not endanger current social security systems and the basis of the welfare State.

---

[139] According to the European Commission only "*24 out of every 1000 female tertiary* graduates" do so in an ICT related subject. Furthermore, only 17% of the ICT specialists in the EU are women. *See*, "International Digital Economy and Society Index 2018, Tech4i2 (Professor Paul Foley, Dr David Sutton, Ian Wiseman, Lawrence Green and Jake Moore) and European Commission, accessed August 10, 2019, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50224; "Women in Digital", European Commission, accessed August 10, 2019, https://ec.europa.eu/digital-single-market/en/women-ict.

[140] You may see below that other instruments like the Villani report treated it in an independent manner.

§ 1.3.12. Environmental Impact

Parliament draws attention to the benefits that AI might have in the fields "*of agriculture, food supply and transport, notably through the reduced size of machinery and the reduced use of fertilizers, energy and water, as well as through precision farming and route optimization"*. In fact, the development of AI can be a game-changer even in the distribution of energy through smart grids which promise to significantly reduce wastefulness. However, Parliament also stresses the AI and AI-equipped robots must be developed in accordance with the European Union's and Member States' environmental policies and should ideally be as "green" as possible by using renewable energies and promoting circular economy in the production chain.

§ 1.3.13. Liability

Parliament draws attention to the fact that the developments in robots and the introduction of AI-enabled features in machines which allow them to make autonomous decisions provide for a shift in the general paradigm in which our liability rules are built. The Institution considers that these machines can no longer be considered simple tools in the hands of human actors such as the manufacturer, the operator, the owner or the user and, thus, new rules and principles may be needed to adapt liability regulation to this new technology.

The Resolution points out a series of shortcomings in the current regulation on product liability in the European Union (please see our chapter on Product Liability, Part II of this Thesis) stating that certain aspects of the current legislation, including the necessity to prove the causal relationship between the machine's action and the damages incurred by the harmed party, may make its applicability highly difficult in the future. According to the EP, the problems are not be contained to product liability, and include other types of liability such as contractual and non-contractual. Furthermore, the Institutions considers that a strict liability regime is needed for manufacturers, operators, owners or users (though more context is not provided).

The EP, calls upon the Commission to submit *"a proposal for a legislative instrument on legal questions related to the development and use of [AI and] robotics and AI foreseeable in the next 10*

to 15 years, combined with non-legislative instruments such as guidelines and codes of conduct". Parliament also asks the EC to make an in-depth evaluation of whether a [purely] strict liability regime should be maintained or a risk management approach focusing not on the person who acted in a negligent manner *"but on the person who is able, under certain circumstances, to minimise risks and deal with negative impacts"* should be applied.

The Document draws attention to the fact that regulating this field adequately will need plenty of preparatory work, including studying present and future human-AI activity. Unlike, what happens in current legislation, Parliament considers that compensation awarded for damages caused by AI-enabled devices cannot be limited by type of damages or extent of damages or forms in which the compensation is offered, in cases other than property damages. The Institution also considers the issue of autonomy, training and self-learning for allocating liability.

Mandatory insurance schemes are presented as a potential solution for the conundrum, along with a fund to ensure compensation in cases where no insurance exists. Parliament calls upon the insurance sector to develop solutions adequate to these new types of machines.

Lastly, Parliament calls upon the Commission to take into account the following issues when designing its future proposal:

"a) establishing a compulsory insurance scheme where relevant and necessary for specific categories of robots whereby, similarly to what already happens with cars, producers, or owners of robots would be required to take out insurance cover for the damage potentially caused by their robots;

b) ensuring that a compensation fund would not only serve the purpose of guaranteeing compensation if the damage caused by a robot was not covered by insurance;

c) allowing the manufacturer, the programmer, the owner or the user to benefit from limited liability if they contribute to a compensation fund, as well as if they jointly take out insurance to guarantee compensation where damage is caused by a robot;

d) deciding whether to create a general fund for all smart autonomous robots or to create an individual fund for each and every robot category, and whether a contribution should be paid as a one-off fee when placing the robot on the market or whether periodic contributions should be paid during the lifetime of the robot;

e) ensuring that the link between a robot and its fund would be made visible by an individual registration number appearing in a specific Union register, which would allow anyone interacting with the

robot to be informed about the nature of the fund, the limits of its liability in case of damage to property, the names and the functions of the contributors and all other relevant details;

f) creating a specific legal status for robots in the long run, so that at least the most sophisticated autonomous robots could be established as having the status of electronic persons responsible for making good any damage they may cause, and possibly applying electronic personality to cases where robots make autonomous decisions or otherwise interact with third parties independently."

With regards to the last issue (legal personality for AI-enabled devices), as we will explain in our chapter about Product Liability (we also touch it in our chapter about the GDPR, namely when studying the problems around identifying the data controller), we do not believe that, as of now, and in our current (or in the short/medium-term) stage of AI-development that it is desirable. However, it does not seem to flow for Parliament's opinion that it supports such a solution. The will of the Parliament appears to be to get ahead of the curve and be prepared for when General AI or sufficiently intelligence precursors to it are developed. Even if it is not something that will happen immediately, there is indeed some value in starting to study it as soon as possible and, therefore, Parliament does not seem to be wrong here[141].

§ 1.3.14. International Aspects

The observations made in this section of the Resolution focus mainly on autonomous vehicles. Parliament considers that there is a need to simplify the framework applicable to traffic accidents based on the Hague Convention of 4 May 1971 on the law applicable to traffic accidents and Regulation Nº 864/2007/EC. Furthermore, it argues that amendments to the referred international Convention and to the Vienna Convention on Road Traffic of 8 November 1968 may be needed to accommodate the challenges brought by this technology. The Resolution also points out that the uniform implementation of international rules must be ensured for autonomous vehicles to work in a seamless manner across the European Union.

[141] It is important to understand that this Resolution's main aim is to make the Commission (and Council) act, by presenting the principal challenges and offering a few (not all) preliminary solutions. It is not supposed to resemble in any manner a final piece of legislation and, thus, general calls to be vigilant regarding a future challenge are not abnormal and should, in fact, be welcomed.

On a more general note, the Resolution calls for strong international cooperation to face the challenges brought by AI in a unified manner (and to avoid a race to the bottom, especially on legal and ethical challenges).

Lastly, attention is drawn to the issue of dual-use items[142] and how the rules governing it, should also be applicable to robotics (and AI).

§ 2. The Commission Starts Building the Foundation for a Specific Legal Framework

§ 2.1. Communication from the European Commission: Artificial Intelligence for Europe

The European Commission's Communication Artificial Intelligence in Europe (hereinafter, AI COM) signals a shift in attitude regarding AI. From a scientific or philosophical perspective, this document does not make for a particularly interesting reading. But the real value of the AI COM is that it reveals the first stages and plans for the implementation of a proper and structured European Strategy for AI.

The first lesson that we should take from the AI COM is that the European Union considers Artificial Intelligence to be one of the key strategic technologies of our century and, as such, that acting on this issue quickly is paramount to its objectives. This is might seem self-evident now, but it was not clear until the release of the AI COM. In abstract, the European Union could decide that its intervention would only be needed when AI achieved the kind of development where establishing rules became necessary to keep freedom of competition and do no more, leaving development, ethical standards, workers' protection, data protection, liability and others to Member States. The chosen path, however, appears to be quite different. In fact, all signs point to a broad scope of intervention by the EU.

The EU wants to help foster the development of AI and, with it, the development of the European Single Market. The AI COM notes the current lack of investment in

[142] Items that can be used for both civil and military purposes. These are particularly important for AI since many AI-enabled robots and algorithms may be used for both civil and military purposes. Drones are a classical example. The same mechanics that potentially allow an autonomous drone to fly and take pictures of nature by itself may also allow it to take the same pictures for intelligence gathering purposes. In fact, most algorithms used to adapt the platforms you most use to your liking (Youtube, Netflix, Spotify etc) could also have some (or a lot of) intelligence gathering value. Of course, there is also the problem of autonomous weapons.

comparison with the US and China, stresses that more public investment from both the Union and Member States in needed and that attracting private investment a must. On the public investment side of the issue, we can point out that in the proposal for the new multiannual financial framework (2021-2027), more precisely in Horizon Europe (the successor to Horizon 2020) the European Commission pledged 7 billion for AI development. The aim is to guarantee, along with Member States and private partners, €20 billion invested in AI until 2020 and at least €20 billion annually from there onwards[143]. Overall one can argue that these objectives are still lacking since they might not be enough to keep with the US and China (you may see the current investment from these countries in their respective chapters).

When the Communication was written, the GDPR was already in force but still not applicable and the wording of the Regulation on a framework for the free flow of non-personal data was not yet finalised. Nevertheless, the European Commission recognised how important data is in the development of AI and choose to make controlling how data is shared one of the key aspects of its AI policy. We will explain below how the early signs point to a significant degree of success on this matter when we analyse the EU legislation on data flow and data protection. One of the key promises of the Communication, a new directive on rules for re-using public sector information (the Open Data and Public Sector Information Directive) has already been fulfilled[144].

The labour market is also one of the main focuses of the EC. On one hand the Commission wants to ensure that the workers' rights are protected while on the other hand, it seems to be aware of the need to assemble a new workforce composed of people who have the skills needed for the development of AI. The reference and use of the expression "no one is left behind" is quite an interesting one. The populist wave that swept across Europe (and also the United States) was, at least, in part motivated by those who were left "behind" by globalisation. This is, even if globalisation and European Integration, generally, were positive for the economy of most countries, some people did not reap its benefits, and in fact, saw their living situations deteriorate. That fact was crucial for the 2016 Brexit Referendum vote, the League and the Five Star Movement election in Italy,

[143] *See*, "Artificial intelligence", European Commission, accessed June 5, 2019, https://ec.europa.eu/commission/news/artificial-intelligence-2018-dec-07_en

[144] Through Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information.

Viktor Orbán's continuous success in Hungary, the Law and Justice Party dominance in Poland and for President Trump's election in the US[145/146].

[145] At the time of writing the United Kingdom is still a member of the European Union, after three extensions to the original date when the country was intended to leave the EU (March 29, 2019 the original, January 31, 2020 according to the last extension). Currently, a general election seems unavoidable and options such as a second referendum or revoking Article 50 altogether are not completely out of the realm of possibility. The same can be said about a no-deal Brexit though. As things stand any prediction of what Brexit's final fate will be should considered now, at least, risky. In what concerns AI, the UK staying in the European Union could be highly beneficial due to the its strong technological and start-up culture and its important economy, amongst others. Nevertheless, we had plenty of opportunities to express our opinions about populism, Brexit and the election of Donald Trump in the past. Therefore, as interesting as an analysis of impact of those occurrences on AI could be, this is not the right place for it and as such we shall guide the reader to those previous works instead of repeating information that would be outside of the scope of this Thesis. *See*, Tiago Sérgio Cabral and Rita de Sousa Costa, *"The European Union's …*: p. 3-15; "A Crise Existencial da União Europeia: Ensaio em torno da realização do projecto europeu no quadro dos desafios geopolíticos e jurídico-institucionais actuais", Tiago Sérgio Cabral and Rita de Sousa Costa, accessed in March 10, 2019, https://institutoeuropeu.eu/images/stories/documentos/Pr%C3%A9mio_Professor_Doutor_Paulo_de_Pitta_e_Cunha/A_Crise_Existencial_da_Uni%C3%A3o_Europeia.pdf; "Homeopathic Democracy: The European Power Struggle over the Spitzenkandidaten", Tiago Sérgio Cabral, accessed in March 3, 2019, https://officialblogofunio.com/2018/03/05/editorial-of-march-2018; "Chronos vs. Brexit: why extending article 50 and delaying Brexit might not be a feasible solution for the EU"; Tiago Sérgio Cabral, accessed in March 3, 2019, https://officialblogofunio.com/2018/12/10/chronos-vs-brexit-why-extending-article-50-and-delaying-brexit-might-not-be-a-feasible-solution-for-the-eu/; "Brexit and the European Football Market: The Consequences for the Premier League and the British Players", Tiago Sérgio Cabral and Rita de Sousa Costa, accessed in November 27, 2018, https://officialblogofunio.com/2016/07/24/brexit-and-the-european-football-market-the-consequences-for-the-premier-league-and-the-british-players/; "Rose-tinted glasses might prove fatal: populists and their performances after the 2017 Dutch general election", Tiago Sérgio Cabral and Rita de Sousa Costa, acesso em November 25, 2018, https://officialblogofunio.com/2017/11/07/rose-tinted-glasses-might-prove-fatal-populists-and-their-performances-after-the-2017-dutch-general-election/; "Democracy, negotiation, personal ambitions and backroom deals: the moment of truth for the spitzenkandidaten", Pedro Madeira Froufe and Tiago Sérgio Cabral, accessed July 2, 2019, https://officialblogofunio.com/2019/07/02/editorial-of-july-2019/.

[146] Numerous academics (and the European Court of Justice) have also expressed their opinion on these issues, we shall point the reader to a brief list, that is no way exhaustive. *See,* Pedro Madeira Froufe, "Editorial of May 2017", accessed in March 27. 2019, https://officialblogofunio.com/2017/05/01/europe-ceci-est-pas-une-pipe/; Pedro Madeira Froufe, "O insustentável peso democrático do populismo: deambulações em torno da União Europeia, de olhos postos em Donald Trump", in *UNIO E-book Volume I: Workshops CEDU 2016*, coord. Alessandra Silveira (Braga: CEDU, 2016): 301-311; "Editorial of July 2016", Alessandra Silveira, accessed in March 30, 2019, https://officialblogofunio.com/2016/06/29/editorial-of-july-2016/; "Editorial August 2016", Katarzyna Gromek-Broc, accessed in February 15, 2019, https://officialblogofunio.com/2016/08/04/editorial-august-2016/; "A Perspective on Brexit", Elaine Dewhurst, accessed February 1, 2019, https://officialblogofunio.com/2016/08/04/a-perspective-on-brexit/; "The voters have spoken. Brexit it is.", Catherine Barnard, accessed in February 1, 2017, https://officialblogofunio.com/2016/08/04/the-voters-have-spoken-brexit-it-is/; "Populist Constitutions – A Contradiction in Terms?", Jan-Werner Müller, accessed in February 1, 2019, http://verfassungsblog.de/populist-constitutions-a-contradiction-in-terms/; Andrew Glencross, *Why the UK Voted for Brexit: David Cameron's Great Miscalculation*, (London: Palgrave Macmillan, 2016), 35ff.; Tim Oliver, "*Fifty Shades of Brexit: Britain's EU Referendum and its Implications for Europe and Britain*", Italian Journal of International Affairs 52,1 (2017): 1-11; John Clarke e Janet Newman, "'People in this country have had enough of experts': Brexit and the paradoxes of populism", Critical Policy Studies 11,1 (2017): 101-116; Sara B. Hobolt, *"The Brexit vote: a divided nation, a divided continent"*, Journal of European Public Policy 23,9 (2016):1259-1277; Alan Ingram, *"Geopolitical events and fascist machines: Trump, Brexit and the deterritorialisation of the West",* Political Geography 57 (2017): 91-93; Anand Menon e Brigid Fowler, *"Hard or Soft? The Politics of Brexit"*, National Institute Economic Review 238 (November 2016): 4-12; Eric A. Posner, "Can it Happen Here?: Donald Trump and the Paradox of Populist Government", in *Chicago Public Law And Legal Theory Working Paper* (n.º 605, Chicago: University of Chicago Law School, 2017); Anna Wyrozumska, "Article 50

Last, but not least, the EC seems aims to strengthen the EU's position as the legal and ethical standard-setter in the world. As previously, legal achievements by the European

[Voluntary Withdrawal from the Union], in *Treaty on European Union (TEU): a Commentary,* eds. Hermann-Josef Blanke e Stelio Mangiameli, (Heidelberg: Springer, 2013), 1385-1418; Afonso Patrão, "Anotação ao artigo 50.º do TUE", in *Tratado de Lisboa*…, 186-189; Alessandra Silveira, "Brexit e o princípio federativo da lealdade europeia: considerações sobre o artigo 50.º do Tratado da União Europeia", in *UNIO E-book*…, 331-348; "Out is out (including in relation to the Mediterranean diet…). On the Article 50 of the European Union Treaty in the light of the federative principle of European loyalty", Alessandra Silveira, accessed February 10, 2019, https://officialblogofunio.com/2016/07/07/out-is-out-including-in-relation-to-the-mediterranean-diet-on-the-article-50-of-the-european-union-treaty-in-the-light-of-the-federative-principle-of-european-loyalty/; "Brexit, The Supreme Court (UK) and the principle of loyalty: on the question of irrevocability of a withdrawal notice", Alessandra Silveira, accessed February 10, 2019, https://officialblogofunio.com/2017/01/26/brexit-the-supreme-court-uk-and-the-principle-of-loyalty-on-the-question-of-irrevocability-of-a-withdrawal-notice/; "R (Miller) v The Secretary of State for Exiting the European Union [2016] EWHC 2768 (Admin) Realpolitik and the Revocation of an Article 50 TEU Notification to Withdraw", John Cotter", accessed February 10, 2019, https://officialblogofunio.com/2016/12/02/r-miller-v-the-secretary-of-state-for-exiting-the-european-union-2016-ewhc-2768-admin-realpolitik-and-the-revocation-of-an-article-50-teu-notification-to-withdraw/; "The right to withdraw the notification to leave the European Union under Article 50 TEU: can we still save the marriage?", Mariana Alvim, accessed February 10, 2019, https://officialblogofunio.com/2017/07/10/the-right-to-withdraw-the-notification-to-leave-the-european-union-under-article-50-teu-can-we-still-save-the-marriage/; "Article 50 TEU: The uses and abuses of the process of withdrawing from the EU", Steve Peers, accessed February 2, 2019, http://eulawanalysis.blogspot.pt/2014/12/article-50-teu-uses-and-abuses-of.html; "Who exactly will 'take back control'? Parliament vs executive after Brexit and the 'Great Repeal Bill'", Steve Peers, accessed February 2, 2019, http://eulawanalysis.blogspot.pt/2016/10/who-exactly-will-take-back-control.html; "Of course you can still turn back! On the revocability of the Article 50 notification and post-truth politics"; Paolo Sandro, accessed February 2, 2019, http://verfassungsblog.de/of-course-you-can-still-turn-back-on-the-revocability-of-the-article-50-notification-and-post-truth-politics/; "After Article 50 and Before Withdrawal: Does Constitutional Theory Require a General Election in the United Kingdom Before Brexit?"; Oliver Garner, accessed February 2, 2019 http://verfassungsblog.de/after-article-50-and-before-withdrawal-does-constitutional-theory-require-a-general-election-in-the-united-kingdom-before-brexit/; "Brexit and the Single Market: You say Article 50, we say Article 127?"; Tobias Lock, accessed February 3, 2017, http://verfassungsblog.de/brexit-and-the-single-market-you-say-article-50-we-say-article-127/; Vernon Bogdanor, "Brexit, the Constitution and the Alternatives", King's Law Journal 27,3 (2016): 314-322; "Reclaiming the Truth: the role of European citizens on countering fake news", Rui Castro Vieira, accessed February 3, 2019, https://officialblogofunio.com/2017/11/29/reclaiming-the-truth-the-role-of-european-citizens-on-countering-fake-news/; Ronald F. Inglehart and Pippa Norris, "*Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash*", Harvard Kennedy School: Faculty Research Working Paper Series (2016): 6; "Behind Trump's victory: Divisions by race, gender, education", Alec Tyson and Shiva Maniam, accessed March 15, 2019, http://www.pewresearch.org/fact-tank/2016/11/09/behind-trumps-victory-divisions-by-race-gender-education; "Clinton Couldn't Win Over White Women", Clare Malone, accessed March, 20 2019, https://fivethirtyeight.com/features/clinton-couldnt-win-over-white-women; Nicholas Confessore and Sarah Cohen, "How Jeb Bush Spent $130 Million Running for President With Nothing to Show for It", *New York Times*, February 22, 2016, accessed March 10, 2019, https://www.nytimes.com/2016/02/23/us/politics/jeb-bush-campaign.html?_r=1; Paola Chavez and Veronica Stracqualursi, "From 'Crooked Hillary' to 'Little Marco,' Donald Trump's Many Nicknames", *ABC News*, May 11, 2016, accessed March 10, 2019, http://abcnews.go.com/Politics/crooked-hillary-marco-donald-trumps-nicknames/story?id=39035114; Cristiano Lima, "Poll: Trump administration edges media in voter trust", Feb. 17, 2017, accessed March 10, 2019, http://www.politico.com/story/2017/02/trump-media-trust-poll-fox-news-235168; Julia R. Azari, *"How the News Media Helped to Nominate Trump. Political Communication"*, Political Communication 33, 4 (2016): 677-680.

Union such as the Charter of Fundamental Rights of the European Union, its standards in product liability, cybersecurity and the GDPR are given as examples of areas where the EU used successfully the weight of its internal market to persuade other countries to follow its standards. The EC quickly points out some areas of AI that need regulatory development explainability, liability; (cyber) security and (algorithmic) bias and discrimination.

We have to point out some merit to this reasoning but, needless to say, also some major flaws. Flexing the single market's muscles indeed works. Hindsight is 20/20 and the EC wrote the AI Communication before the GDPR was applicable, but we have the luxury of writing after, so we cannot avoid its example. Countries like Israel, New Zealand, Argentina, Colombia, South-Korea, Brazil and Japan made or are making developments to their legislation to mirror the EU's own[147]. In Japan's case, data protection negotiations were part of the negotiations for the free trade deal with the country[148]. Even in the US states such as California[149] already implemented data protection legislation[150] (albeit significantly less strict than the GDPR) and, in an interesting twist, big tech companies like Google, Amazon and Apple are asking Congress to pass legislation on data protection[151].

That notwithstanding, to flex your trade muscles, you have to have trade muscles. Therefore, the EU depends on the continuing success of its single market to act as the regulatory standard-setter in AI. Thus, the EU needs to develop its AI market, but not exclusively as user, also as a producer. To do so the EU needs a flexible and business-

---

147 *See*, "Europe's new data protection rules export privacy standards worldwide", Mark Scott and Laurens Cerulus, accessed April 10, 2019, https://www.politico.eu/article/europe-data-protection-privacy-standards-gdpr-general-protection-data-regulation/; "Communication from the Commission to the European Parliament and Council: Data protection rules as a trust-enabler in the EU and beyond – taking stock", European Commission, accessed September 10, 2019, https://ec.europa.eu/commission/sites/beta-political/files/communication_from_the_commission_to_the_european_parliament_and_the_council.pdf.

148 *See, "*Commission Implementing Decision of 23.1.2019 pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council on the adequate protection of personal data by Japan under the Act on the Protection of Personal Information", European Commission, accessed April 10, 2019, https://ec.europa.eu/info/sites/info/files/draft_adequacy_decision.pdf

149 The California Consumer Privacy Act of 2018 becomes effective on 1 January 2020. One should note that, specifically regarding data breaches, 50 states in the US, the District of Columbia, Guam, Puerto Rico and the Virgin Islands already have rules regarding the notification to the affected parties. *See*, "Security Breach Notification Laws", National Conference of State Legislatures, accessed October 10, 2019, http://www.ncsl.org/research/telecommunications-and-information-technology/security-breach-notification-laws.aspx.

150 *See,* "The California Consumer Privacy Act", accessed October 10, 2019, http://www.mondaq.com/unitedstates/x/750962/Data+Protection+Privacy/California+Consumer+Privacy+Act+of+2018+Full+Text

151 *See*, "Silicon Valley finally pushes for data privacy laws at Senate hearing", Dan Tynan, accessed April 5, 2019, https://www.theguardian.com/technology/2018/sep/26/silicon-valley-senate-commerce-committee-data-privacy-regulation;"Open Letter on Privacy", VV.AA., accessed September 10, 2019, https://s3.amazonaws.com/brt.org/BRT-CEOLetteronPrivacy-2.pdf.

friendly legal framework that does not jeopardise its citizens' fundamental rights. It is a delicate balancing act.

In addition to harmonization, cooperation is key and, knowing this fact, the Commission stresses the need to ensure an adequate level of cooperation both between the European Member States with initiatives such as the Coordinate Plan on AI[152] , also between the Union and Member States and the relevant stakeholders with initiatives like the creation of the High-Level Expert Group on Artificial Intelligence and European AI Alliance.

The Commission further calls for global solutions to address issues such as the effects of AI on the environment, competition in AI and the military use of AI.

An interesting point, that is certainly worth pointing out, is that the Commission takes a very wise position on monitoring the development of AI. First, clarifying that the current debate is frequently based on non-factual assumptions and that continuous benchmarking of "*the technical capabilities of AI components and systems to give a realistic understanding of where the technology stands, and help increase public awareness*" is needed to assess when to propose the necessary regulation. Of course, this is easier said than done, because too early and you are going to be discussing something for which you will not have all the necessary information and potentially create bad regulation and damage the market, but too late might mean having unregulated AI creating dangers to the fundamental rights of European Citizens.

[152] Building on the Declaration of Cooperation on Artificial Intelligence, previously signed by all Member States, plus Norway. *See*, "Declaration of Cooperation on AI – Signed by Austria, Belgium, Bulgaria, Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, UK and Norway", EU Member States, accessed March 5, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50951. Croatia, Cyprus, Greece and Romania joined after. *See*, "EU Member States sign up to cooperate on Artificial Intelligence", European Institutions, accessed March 5, 2019, https://ec.europa.eu/digital-single-market/en/news/eu-member-states-sign-cooperate-artificial-intelligence.

§ 2.1.1. Commission Staff Working Document: Liability for Emerging Digital Technologies[153]

Considering that the problem of liability should be addressed separately, and in more depth, the European Commission released an accompanying document on the issue of liability for emerging digital technologies.

The Working Document is not specifically direct towards Artificial Intelligence, covering other emerging technologies such as the Internet of Things (IoT) and advanced robotics and autonomous systems. More likely than not, advanced robots and autonomous systems will be AI-enabled and frequently use Machine Learning algorithms and models. Many devices that are part of IoT are also part of the AI and Machine Learning family. Your smartphone, your smart speaker and your smart house are all IoT devices, but they would not be able to perform a relevant number of the tasks we currently expect from them without advanced Machine Learning algorithms and robust quantities of data.

While conceding that allocating liability for damages caused by emerging technologies is a difficult task, the Commission stresses that an adequate approach is deeply needed. The implemented solution must protect the users and give them the peace of mind of knowing that these technologies will rarely fail and that when they do the legal framework possesses the needed remedies to ensure that they are adequately compensated for any damage that was caused to them. Meanwhile innovative companies from and/or working in the EU need the legal certainty of a framework that is updated to current times, adapted to their needs and to the specific characteristics of emerging technologies, realistic regarding the unpredictability and current state of technological development and that does not hurt the competitiveness of European companies in the global market.

After a brief exposition of the European Legal instruments containing safety rules that may be applicable to AI and other emerging technologies and of the standardisation efforts being enacted by European Standardization Organizations, the Working Document describes, shortly, the existent liability regimes in the European Union and Member States relevant for emerging technologies. Both the issues of safety in AI and standardisation, when relevant, shall be properly analysed below. In fact, the Working Document offers no

153 An accompanying document to the Artificial Intelligence for Europe Communication. *See*, "European Commission Staff Working Document: Liability for emerging digital technologies. Accompanying the document Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial intelligence for Europe", European Commission, access December 10, 2018, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51633.

more than a brief description, thus not entailing a dedicated analysis. The Working Documents also bring nothing to the table on the examination of the general principles of liability, again it is just a (very brief) description of the Product Liability Directive and of what is fault-based non-contractual liability and strict liability. There are some interesting questions regarding *a)* the difficulty in predicting (and tracing the path that led to a certain) behaviour of AI-enable technologies which can create a challenge in allocating liability; *b)* the issue of bad or corrupted data*; c)* the fact that certain AI-enabled technologies can be seen as providing a service and thus not fall within the scope of the Product Liability Directive and; *d)* the issue of software updates and upgrades. All the issues mentioned shall be properly examined by us in our Chapter on Product Liability.

Four different case studies are analysed within the context of the Working Document. They are: *i)* autonomous drones; *ii)* autonomous cars; *iii)* smart ecosystems and; *iv)* the issue of cyber-attacks against IoT devices. While the analysis is indeed quite interesting and its reading is wholeheartedly recommended, we shall not delve deeply into the abovementioned section for questions related to both time and scope of this Thesis.

Wisely, the Working Document recommends further engagement with these issues and deeper analysis to assess whether concepts and elements currently contained within the EU/national liability framework should be revised in accordance with new challenges arising from emerging technologies. The document asks whether a liability of the guardian regime, similar to what exists for animals, should be considered for AI. It argues that while it is true that AI is autonomous and, sometimes, unpredictable, the same could also be said for animals. In fact, one may even argue that AI's unpredictability can be "minimised" through programming, something that is not true for animals. While this as a mere question intended to do no more than to stir the legal discussion around the issue, there are more than a few shortcomings to the analogy. First, the scope and complexity of the tasks where humans will be supported by AI or where AI will work independently for humans is, arguably, much broader than animals. Your dog may guard your house and your sheep, a horse may sometimes transport you etc. however, AI will (help) decide if you are the most qualified candidate for a job opening, by analysing thousands of parameters of which the programmers themselves may not be necessarily fully aware. AI will drive cars at high speeds in busy highways and cities, recognise and charter the most adequate path to your destination. AI will replace human workers in a plethora of (mostly) repetitive but still sophisticated and human-level tasks, the nature of which would be very difficult for an animal to understand. Further examples could be provided, but it would be redundant for

our argument. Still, we must note that, in addition to this fact, understanding AI for the general end-user will likely be more complicated than understanding an animal for the person who uses animals to collaborate in its work. Therefore, while the idea of guardian liability may be acceptable in certain circumstances for AI, the analogy with the animal-case is not a strong one.

The difficulties of establishing adequate liability procedures for different AI-based systems are also drawn to attention. Deciding which systems should be fault-based and which should be based on strict liability or other possibilities is a significant challenge. Questions related to the burden of proof, type of damages that are compensable and right to redress between different actors in the value chain are also addressed as is possible liability for cyber-attacks (see below, our Chapter on Product Liability for more information).

§ 2.2. Communication from the European Commission: Coordinated Plan on Artificial Intelligence[154]

This Communication serves as an introduction to the Coordinated on Artificial Intelligence (hereinafter, "the Coordinated Plan") previously agreed with the Member States and contained within its annex. The Communication itself is used to highlight the main objectives and initiatives in the Coordinated Plan. Our analysis shall cover both the documents concurrently.

As an introductory statement, the Commission states that 5 Member States had already answered the challenge and adopted national AI strategies with a dedicated budget. Nevertheless, since that number is still very much a minority, it encourages the remaining Member States to adopt national AI-strategies by mid-2019. It is also noted by the Commission that it and Member States will agree on indicators for monitoring AI

---

154 *See,* "Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions - Coordinated Plan on Artificial Intelligence", European Commission, accessed January 15, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56018; "Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions - Coordinated Plan on Artificial", European Commission – Annex", accessed January 15, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56017.
The Coordinate Plan follows the Declaration of cooperation on Artificial Intelligence, where Member States had already declared their will to cooperate on the development of AI in the future. *See*, "Declaration of Cooperation on AI", European Commission, accessed February 10, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50951.

development and uptake across the European and to Union assess the efficacy of the strategies implemented.

Currently, the European Union harbours 24% of the players in the AI industry, a little more than the 23% of China and a few below the 25% of the US. However, if we examine the number of players per capita, the distance between the US and the EU widens, and the EU lags behind countries like Israel, Canada, Singapore and Switzerland. In research, 30% of papers submitted to the top AI international conferences come from the European Union. However, as the Communication explains, the EU is strongly lagging when it comes to investment being a distant third with $3 to $4 billion in 2016, against the $8 to $12 billion invested in Asia and the $15 to $23 billion in North America[155]. Thereby, the European Union needs to ensure a boost in investment until its value achieves, at least, €20 billion a year. Horizon 2020, already increased the funds invested in AI, making available €1.5 billion for the period 2018-2020. In the proposal for a new Multiannual Financial Framework (2021-2027) the European Commission proposed, through Horizon Europe (the successor of Horizon 2020) and the Digital Europe Programme, annual investments of, at least, €1 billion per year in AI. To ensure that the goal of €20 billion a year is achieved, the public sector will have to invest €7 billion yearly, with the rest being provided by the private sector.

Taking this in mind, the Commission considers that coordinated public efforts between the Union and Member States are needed as coordinate action will leverage more private investment. Furthermore, ensuring that there are no barriers to trade is key to allow European business to succeed, scale-up, internationalise and invest.

§ 2.2.1. Cooperation between the EU, Member States and Stakeholders

The Commission stresses the importance of reinforcing cooperation with the private sector and fostering initiatives to provide funding (both from public and private sources) to companies scaling-up and start-ups developing solutions based on emerging technologies. Particular attention should be given to AI technologies which are, by design[156], in accordance with European standards and policies, ethical and secure. The

[155] *See*, "10 imperatives for Europe in the age of AI and automation", McKinsey Global Institute, accessed March 20, 2019, https://www.mckinsey.com/featured-insights/europe/ten-imperatives-for-europe-in-the-age-of-ai-and-automation.

[156] In the Coordinated Plan ethics by design is defined as development where *"ethical and legal principles, on the basis of the General Data Protection Regulation, competition law compliance, absence of data bias are implemented since the*

Commission also draws attention to the steps taken for the implementation of the European Innovation Council (hereinafter, "EIC"). As of this writing, 22 stakeholders were appointed to oversee the rollout of the pilot for the EIC. It is expected that the EIC will be fully functional in 2021 with Horizon Europe[157]. Current tools available for AI funding are listed along with advice on prioritizing activities in the following manner: *"i) financing a portfolio of innovative AI/blockchain companies; ii) developing a dynamic EU-wide investors community focusing on AI; iii) multiplying investments at the national level by involving the national promotional banks that are willing to participate; iv) incentivising private sector investments and v) making Europe become more attractive for start-ups to stay and grow. In the following years".*

Plans are outlined to foster cooperation between research centres working on AI across the European Union. Cohesion funds are also pointed as a manner of ensuring converge of AI uptake across Member States and sectors. Guaranteeing that 5G networks are available across the EU and investing in standardisation are also considered essential to integrate AI in SMEs.

In the Coordinated Plan, making European public administrations the frontrunners in the use of AI is also nominated as an objective.

The Commission calls upon Member States to *"speak with one voice"* in the arena of international politics, in the matter of AI Regulation. In this manner, the Commission hopes to promote the European AI vision across the world (maybe mirroring the success it had on data protection).

§ 2.2.2. Regulatory Initiatives

Completing the Digital Single Market is considered key to achieving the EU's objectives as a leader in AI.

Making secure, robust quality data available for a *"broad range of users across borders is a cornerstone of European policy"*. Regulatory measures such as the non-personal data regulation and the new Open Data and Public Sector Information Directive are critical for

beginning of the design process" and security by design development where *"cybersecurity, the protection of victims and the facilitation of law enforcement activities should be taken into account from the beginning of the design process".*

[157] "Appointment of members of the High-Level Expert Group on the Impact of the Digital Transformation on EU Labour Markets", European Commission, accessed July 17, 2019, https://ec.europa.eu/digital-single-market/en/news/appointment-members-high-level-expert-group-impact-digital-transformation-eu-labour-markets.

that[158]. Securing the necessary computing power for processing large quantities of data is also addressed through the European High-Performance Computing Joint Undertaking[159]. The EU also seems intent to include rules on data processing on future trade agreements with its partners.

There are no significant changes in the Commission's position regarding the ethical implementation of AI and legislative initiatives needed for securing this objective. The ongoing assessment of the adequacy of European and Member States' liability legal framework is also referred. Interoperability and the promotion of *"open, fair, machine readable, standardised and documented"* formats are also signaled as priorities.

Methods of policy experimentation like regulatory sandboxes and innovation deals can be used to assess whether the currently implemented rules are adequate for the development of artificial intelligence or if amendments should be considered.

Close attention should be paid to the impact of AI in areas like competition, consumer protection, intellectual property and the impact of AI in human behaviour.

§ 2.2.3. Sectorial Concerns

The labour market appears as one of the areas where the Commission feels that a deeper intervention is needed to train new professionals with the skills necessary to work with emerging technologies, to retain those professionals in the face of lucrative offers frequently coming from abroad and to attract talent from abroad to reinforce the European workforce. Currently, most Member States are facing shortages of ICT professionals and the current formative offer is not enough and not evenly distributed.

The European Commission also calls upon policymakers to develop solutions to allow for inclusive and seamless transition into the new digital labour market. The

---

158 The Commission's Expert Group for the Observatory of the Online Platform Economy is tasked with *"exploring issues in AI-related regulatory areas, such as data access, online advertising, online advertising and the role of algorithms in the digital platform economy". See,* "Expert group to the EU Observatory on the Online Platform Economy", accessed in October 13, 2019, https://ec.europa.eu/digital-single-market/en/expert-group-eu-observatory-online-platform-economy.

159 This initiative aims to install in the EU *"two supercomputers that will be among the top 5 in the world and at least two other that would today rank in the global top 25 for Europe's private and public users scientific and industrial users, for use in more than 800 scientific and industrial application fields"* and *"developing a European supercomputing ecosystem, stimulating a technology supply industry, and making supercomputing resources in many application areas available to a large number of public and private users, including small and medium-sized enterprises". See, "*The European High-Performance Computing Joint Undertaking – EuroHPC", European Commission, accessed July 17, 2019, https://ec.europa.eu/digital-single-market/en/eurohpc-joint-undertaking.

document also contains a promise to create a High-Level Expert Group on the Impact of the Digital Transformation on EU Labour Markets. Said promise is fulfilled and the report about the impact of the digital transformation on EU labour markets was made public in April 2019.

Lastly, security concerns regarding the use of AI (including potentially on automatic weapons) and the protection of AI against external attacks are very briefly addressed.

§ 2.3. Ethics Guidelines for Trustworthy AI

The first of two deliverables by the HLG builds the concept of "Trustworthy AI". Trustworthy AI is the logical evolution of the concept of ethics by design and for an AI to be "trustworthy" it has to fulfil three requisites: *a)* "be lawful, complying with all applicable laws and regulations"; *b)* "be ethical, ensuring adherence to ethical principles and values"; and *c)* "be robust, both from a technical and social perspective", since, even with good intentions, AI systems can cause unintentional harm. The ethics guidelines do not address the first challenge, making them less appealing in our scope than the second deliverable, the Policy and Investment Recommendations for Trustworthy AI. However, it is highly likely that we see some of the ethical recommendations contained within the Ethics Guidelines eventually enshrined into law, and thus an analysis of the main points should be performed.

One preliminary note, even though the Guidelines are obviously not binding, they do set themselves as a method to which stakeholders interested in showing their compromise to Trustworthy AI can adhere voluntarily and in their current form.

While the Guidelines do not specifically address the legal component of AI regulation, their ethical analysis of rules that are already fundamental rights in EU and Member State law offer an interesting insight on how those principles can be interpreted as to adapt to the new AI-enabled world or can be enshrined in legal diplomas dealing with this question.

The principle of human dignity must be used to protect humans against what we can call objectification of the human being by AI. The HLG considers that this includes being "*sifted, sorted, scored, herded, conditioned or manipulated*". The question of individual scoring by AI-means frequently appears in the deliverables of the HLG and should be

understood as a complete rejection of a China-like paradigm where citizens could be given scores and can see their rights culled due to low results[160]. It is our opinion that such a measure would not be possible at all in the European Union, as it would infringe the principle of human dignity, even without any updated interpretation for AI, and relatively influent legislation such as the GDPR. Still, it is a valid concern and we shall deal with it further in the context of this work.

The Principle of Individual Freedom should be interpreted in a manner that denies AI-enabled mass control surveillance as well as manipulation, amongst others. Bottom line AI should not be used to infringe upon the free will of human beings (as far as that is possible), it should instead complement them.

Respect for democracy, justice and the rule of law addresses some important concerns with AI possibly undermining the democratic process or being used to do so or used in a manner that harms due processed and the right to a fair trial. This reflects certain cybersecurity concerns that bots may be used to spread fake news and influence democracy or even straight-up hack elections. Furthermore, use of AI in a pre-sentencing or even sentencing context is already being implemented in other jurisdictions such as the US (with very mixed results as you can see in our Chapter about the GDPR) and the HLG appears to reject this notion.

Equality, non-discrimination and solidarity in the context of AI should, in accordance with the HLG, follow a logic of non-bias and accessibility (to vulnerable groups). To this we would add the equal access to the opportunities provided by AI regardless of income.

[Respect for] Citizens' Rights is analysed in a context of positive rights, including *"the right to vote, the right to good administration or access to public documents, and the right to petition the administration"*. The HLG states that AI should foster and not hinder the exercise of these rights by the citizens. We would add that negative rights are equally important in this context and the European Union should support and ensure that AI is not used in a manner that diminishes their exercise (in line with the remaining principles). The HLG draws attention to the fact that third-country citizens should not be forgotten, even if they are not EU nationals. To this consideration we add (and will develop below) that the EU should resist populist urges to, for example, try to identify illegal immigrants through the

[160] See our chapter on China.

use of AI. In fact, future European legislation on AI should contain the humanistic standards expected of the bloc and apply proportionate means to achieve its objectives.

These principles can and should be connected to ethical principles in the context of AI systems developed by HLG and equally to requirements for implementation of Trustworthy AI.

Respect for human autonomy should be tied to the principle of individual freedom and also to the respect for citizens' rights. It includes the requirements of non-objectification and complementariness. Human oversight and human agency (requirements for achieving a Trustworthy AI) should also be considered to be under the principle of individual freedom and required by the respect for human autonomy.

Human agency is understood by the HLG as meaning the AI should act as a human agent instead of a replacement. AI should help humans in making their decision. The HLG quotes Article 22 of the GDPR, which we have deeply analysed in our chapter regarding the GDPR to which we refer. Human oversight means that, in the end, a human should, at least have the capability to override the AI's decision. Further, the decision about when and to what purposes use AI-enabled systems should still rest with human actors. Explicability (explainability) should also be considered to be under human autonomy and the right to self-determination. The HLG deals with explicability both under explicability as an ethical principle and under transparency as a requirement for Trustworthy AI. It defends an adapted level of explicability in accordance to the impact of the AI's decision on the person. For decisions that have a significant impact on someone's life a full explanation is needed of the system's decision-making process. The degree to which the AI input is taken into consideration into the final decision should also be disclosed[161]. Humans should also be made aware that they are contacting with a computer program and not with a person, and be able to ask for human intervention if they feel the need to do so (communicability).

Within the (legal) principle of human dignity, we should include the principle of prevention of harm. The HLG considers that the principle goes further than merely avoiding mental or physical damage caused by AI-enabled systems. It should also not work to exacerbate asymmetries of power or information. In our opinion this should be

[161] On this issue the HLG position is in line with what we defend regarding the right to explainability under the GDPR. However, as we will further see below the HLG actually goes further than what is possible to achieve under the GDPR, asking for explanation of decisions that are not based solely on automated decision-making or in personal data. We will further analyse its proposals in our conclusions/policy recommendations.

considered a principle of prevention and compensation of harm, as legislation is needed to ensure (as admitted by the HLG in the Policy and Investment Recommendations for Trustworthy) that when harm does arise from the use of AI, remedies for compensation are in place.

Fairness is almost transversal to all the abovementioned legal principles. In fact, in any legal text to be produced regarding AI, fairness should be considered independently as a legal principle. We can see that the HLG drew inspiration from fairness in data protection by requesting the possibility of identification of the entity responsible for the use of AI, explicability of decisions and right to context decisions by AI systems. Fairness also contains non-discrimination, individual freedom and equality of opportunities aspects to it. Quality and integrity of data are also key measures to ensure fair and non-discriminatory AI development as is accountability, including auditability and redress.

Regarding the technical and non-technical methods proposed by the HLG to achieve Trustworthy AI, we would like to highlight the support for the ethics by design approach that, in fact, should be called Trustworthiness by design, since this is a broader concept and the need for investment in explainable AI (which we will address below, namely in our chapter about the GDPR). For non-technical methods, we feel like Regulation in a more high-level manner and soft law or standardisation in a more technical and low-level manner will be key for the development of Trustworthy AI, as will certification to make "EU-approved" AI easily identifiable to consumers and buyers in general.

Furthermore, we would like to add that the idea of Trustworthy AI can only be realised with Trustworthy Regulation of AI. That is, ensure that citizens are as involved as possible, directly and through their representatives. For this, a transparent ordinary legislative procedure is key, along with actively divulging and promptly making available documents from trilogue negotiations to allow common citizens to lobby for their preferred solutions. Of course, this along with the general methods of engaging the public, including events, public consultations, media campaigns amongst others. It is very important that the European people understand the added value of EU-wide regulation for AI, that they will be able to influence the process and that they can trust the EU to deliver as only then will they trust and prefer any AI-enabled devices that comply with said regulatory approach.

Lastly, it is important to point out the Ethics Guidelines developed an assessment for Trustworthy AI which, even if it already in a pilot stage, may be of some use for manufacturers and developers.

§ 2.3.1. Communication from the European Commission: Building Trust in Human-Centric Artificial Intelligence

This Communication, published by the EC shortly after the Ethics Guidelines were made public, does not offer much in terms of content. In fact, it pretty much restricts itself to repeating a summarised version of some of the Ethics Guidelines requirements and endorsing them. It reveals that a new and updated version of the Assessment for Trustworthy AI will be released by the HLG at the beginning of 2020, with feedback from relevant stakeholders, when the Commission will also evaluate the progress and proposed any necessary new steps. This timeline is sensible because it is shortly after the new Commission takes office. The EC also promises (again) to: *i)* launch *"a set of networks of AI research excellence centres through Horizon 2020"*; *ii)* set up *"networks of digital innovation hubs focussing on AI in manufacturing and on big data"* and; *iii)* to start *"start preparatory discussions to develop and implement a model for data sharing and making best use of common data spaces"*.

Lastly, the Commission will endeavour to make Trustworthy AI as an International Standard through cooperation with partners in third countries.

§ 2.4. Definition of AI HLG

In conjunction with the Ethics Guidelines on Trustworthy AI, the HLG issued a report centring on the issue of the adequate definition of AI to be used on its deliverables (and, potentially, in the future in any legislative proposal to be submitted by the Commission). The report contains a bit more than a simple proposal on a definition for AI though, as it also includes a description of AI's capabilities and research area. However, by the HLG's own admission this no more than a very crude oversimplification of the state-of-the-art. With this in mind, we shall not consider that part of the document.

In regards to the proposal of a definition itself, we refer to our chapter about the definition of AI where it was already analysed.

§ 2.5. Policy and Investment Recommendations for Trustworthy AI

The second deliverable by the HLG builds upon the more general considerations of the Ethics Guidelines by proposing specific measures to be implemented across the EU. From fostering investment and public/private sector collaboration to data, the main points of action are similar to the ones in the Ethics Guidelines. To avoid a repetition of our considerations in the section regarding the HLG's Ethics Guidelines, we shall focus on two main points: *i)* the HLG's recommendations for establishing an appropriate governance and regulatory framework, since the question was not dealt with within the Ethics Guidelines and; *ii)* some key sectorial measures that we believe could be especially important in the development of AI across the EU and that have not been deeply analysed in other sections.

Regarding point *ii)*, it is quite interesting to see how the HLG is drawing inspiration from some Member States, such as Finland, in proposing the creation and promotion of AI-related online courses (as an example, massive open online courses). As noted by the HLG, ensuring universal access to this type o recourse is key to success, and may we add a difficult challenge. The idea for a European AI Awareness Day, if properly organised and with the necessary resources, may contribute to familiarise European citizens with the technology and to reduce mistrust.

Mandatory self-identification for AI systems in cases where, in the context of an interaction between a human and AI and when the human can reasonably mistake the AI for another human, should be a requirement included in any legislative intervention of the EU in this area.

It is also interesting to note how the HLG rejects AI-based mass surveillance. The European Union may not (directly) have the competences do ensure a ban on this type of technology though. Therefore, a dual approach could be considered where the EU cooperates with Member States and partners to achieve this result.

The HLG proposes introducing a duty of care for developers of consumer-oriented AI systems to ensure that these can be used by all intended users, fostering a universal design approach. This proposal should be approached very carefully. This is the type of proposition that must be painstakingly planned to ensure that the final process achieves the adequate level of prowess. Therefore, we must take into account that, on one hand, the principle of non-discrimination should already provide for non-discriminatory

and non-biased AI. On the other hand, this type of measures should not be used by the State to shift the obligation of taking care of its most vulnerable citizen's to private entities.

Additionally, AI-enabled devices is an umbrella term, that encompasses a plethora of equipment, whose requirements for being accessible and importance of ensuring said accessibility will vary significantly, potentially creating a difficult challenge to overcome. However, taking into account the importance of making sure no one is left behind by the AI-revolution,[162] some measures can and should be implemented. First, the EU and Member States should support research and development of AI-enabled accessible devices. Second, the Proposal for a Directive on the approximation of the laws, regulations and administrative provisions of the Member States as regards the accessibility requirements for products and services, could be revisited in the future to integrate it with AI-enabled devices. Since there is a political agreement on the proposal and it was adopted in the EP, with only the Council's formal adoption missing, the Directive itself could be amended in the future. We should note, however, two things. First, a Regulation seems like the most appropriate legislative act for this type of measure, as it seems like the most adequate means to ensure the protection of vulnerable citizens (avoiding any confusions that might, and generally, do arise in the implementation/transposition), and provides a higher degree of security to economic operators for similar reasons. A maximum harmonisation Directive may also work, but we see no advantage in this solution as opposed to a Regulation, and there are quite a few disadvantages (again, even in maximum harmonisation Directives, States sometimes take excessive "creative" liberties in implementation). Second, the current proposal already includes some important measures that would also be applicable to key types of AI-enabled devices such as the ones applicable to "general purpose computer hardware and operating systems" and to "audiovisual media services and the related consumer equipment with advanced computing capability"[163]. Still on this subject, we would note that the fact that measures arising from the Directive will only be applicable 6 years after the Directive enters into force may leave citizens with an inadequate level of protection during that time (for no reason, since the measures contained in the Directive are not new, just vary between Member States) and economic

[162] And thus, managing to protect both fundamental rights and avoid rising populist sentiments due to the economic inequality arising from AI-deployment and use.

[163] *See*, "European Accessibility Act", European Commission, accessed June 20, 2019, https://ec.europa.eu/social/main.jsp?catId=1202 ; "European Accessibility Act: final steps on the European level – first steps on the national level", European Union of the Deaf, accessed June 20, 2019, https://www.eud.eu/news/european-accessibility-act-final-steps-european-level-first-steps-national-level/.

operators with an incoherent and fragmented set of rules during a key moment for technological development.

Investment, partnerships between public and private entities are also addressed, in line with the explained above. Making sure that SME's have the technical and legal expertise available to introduce and invest in AI solutions to enhance their business practices.

The need for the right to request human intervention when there is a significant impact on the individual is reinforced. The HLG suggests a more extensive formulation than what is contained in the GDPR (because it would also be applicable when there is no personal data processing). A set of measures to ensure digitalisation and modernisation of public administration is suggested to ensure it is faster, more accessible and easier to use (see below our Section about Estonia). To realise this objective universal and high-quality internet access is required, in accordance with the HLG.

To provide the necessary data for AI development in the European Union, the HLG proposes measures like the transformation of public data into digital format, the establishment of data-sharing infrastructures and creation of a European centralised annotated public non-personal databases that are accessible to legitimate stakeholders for algorithmic training[164]. This measure would foster the development of standards and

---

164 One other interesting measure is the possible creation of "data trusts" for specific sectors. Generally, a Trust is a legal relationship where the trustee, is given by the trustor an asset to hold/manage, for the benefit of a third party (the beneficiary). To be applicable in this context, the data subject would have to be both trustor and beneficiary, therefore the contract would look more like an authorization to manage a pool data than a Trust as we are generally aware of it. This was suggested, for example, by the growing the artificial intelligence industry in the UK report commissioned by the UK Government. The report states that: *"to use data for AI in a specific area, data holders and users currently come together, on a case by case basis, to agree terms that meet their mutual needs and interests. To enable this to be done more easily and frequently, it is proposed to develop terms and mechanisms for these parties to form, between them, individual "data trusts" to enable AI to be developed to meet the needs of the parties involved and allow data transactions to proceed with confidence and trust. These trusts are not a legal entity or institution, but rather a set of relationships underpinned by a repeatable framework, compliant with parties' obligations, to share data in a fair, safe and equitable way".*
We do not see such a measure as impossible, but it certainly brings some problems regarding the provision of information to the data subject and the legal basis for the processing of personal data. While some options such as establishment of multiple contracts between data subjects and an entity acting as a Trustor, or even basing the processing on the legitimate interests of the data controller or even public interest, the stronger legal basis would be the creation of specific legislation regulating the creation of management entity in the sectors, what data could be manage, if profit arose how it would be distributed, etc. To ensure that said legislation is compatible with the GDPR, and that more data is available it would make sense to establish the mechanism at a European level. *See*, "Growing the artificial intelligence industry in the UK, Dame Wendy Hall and Jérôme Pesenti, accessed June 5, 2019, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf; "Data trusts: why we are interested", Jack Hardinges, accessed June 7, 2019, https://theodi.org/article/what-is-a-data-trust/#1527168650599-ae3e3b8c-e22a62d2-2d92; "Defining a 'data trust'", Jack Hardinges and Peter Wells, accessed June 7, 2019, https://theodi.org/article/defining-a-data-trust/.

interoperability and ensure that errors in data could be corrected in one centralised database. Data donation schemes where data subjects would be able to "donate" their data to a specific cause are deemed a solution that could have merit. Procurement contracts could also include clauses where the data generated within the performance of the contract and that does not infringe IP rights is made available to the public. Making access to data on fair, reasonable, and non-discriminatory terms (hereinafter, "FRAN"), in line with the regime for key patents and incentivizing or even requiring mandatory interoperability between key market players, could be a competition-law inspired solutions that would allow new market players or market players with access to less data to develop cutting-edge AI. Still, the definition of what would be FRAN for data would be quite troublesome. It could also be counterproductive, as large market players could suddenly license interoperable databases from each other and widen the gap with smaller companies who would be unable to acquire the same variety of licenses.

Furthermore, for personal data, it could ultimately turn it into a commodity to be bought, sold and licensed at will. While this is not impossible under the GDPR, it could seriously hinder the exercise of rights by data subjects and finding a legal basis for licensing in this context may prove difficult. We should not forget that patents will not (likely) disappear all of a sudden, but if you obtain a license a database with personal data from one million data subjects and then after providing them with the adequate information, all of them exercise their right to erasure or object to the processing of their personal data, you may have paid for a database that is now completely worthless.

The HLG also argues for a general explanation right for AI-informed decisions taken by public administration and a ban in mass scoring of individuals.

Investment in both research projects and the necessary infrastructure for AI deployment and development (such as 5G networks) are also indicated as necessary priorities for the European Union.

Interestingly, Jack M. Balkin considers that companies like Facebook or Google should be considered Information Fiduciaries and have a duty of care regarding the information they collect about their users. However, Balkin's theory, while interesting in the American context, holds less value in the European context, where the GDPR provides for broader rights for the data subject and obligations for data controllers than what would be possible under the information fiduciary theory. Any development regarding data trusts must happen within the context of the GDPR's principles. *See*, Jack M. Balkin, "Information Fiduciaries and the First Amendment", UC Davis Law Review 49,4 (2016): 1183-1234.

§ 2.5.1. Establishing an Appropriate Governance and Regulatory Framework

The HLG requests that the EU adopts a principle-based and risk-based approach to regulating AI. Instead of laying down excessively detailed legislation, that may easily become outdate due to rapid advancements in technology, the HLG feels that creating high-level principles based on "AI and EU values" should be the solution pursued. Furthermore, a risk-based approach means that *"higher the impact and/or probability of an AI-created risk, the stronger the appropriate regulatory response should be."*. The HLG considers that concepts as an unacceptable risk should be defined by community at large, still we must point that, written between the lines, the HLG appears to pass the message that mass scoring and mass surveillance of individuals would clearly be considered as unacceptable. Furthermore, the Policy and Investment Recommendations call upon the Commission to transform the idea of *"Trustworthy AI into a concrete set of indicators that can be used for monitoring the convergence of the European market towards the desired policy goals"*. According to the HLG, legislation must be adapted to by context and sector. In fact, it goes a bit further when considering that AI systems deployed by the private sector that have the potential to have a significant impact on human lives, for example by interfering with an individual's fundamental rights at any stage of the AI system's life cycle[165] could be required to perform a Trustworthy AI assessment, a relevant stakeholder consultation and employ adequate safeguards before entering into the market.

A general assessment of existing EU laws that may affect AI systems is requested in the Recommendation to ensure coherent solutions. Proposals are then divided by areas of the law.

a) For civil liability and accountability rules the HLG first asks if whether *"it is necessary or desirable to introduce traceability and reporting requirements for AI applications to facilitate their auditability, ex-ante external oversight before AI systems can be deployed, systematic monitoring and oversight by competent authorities on an ongoing basis, and the obligation for meaningful human intervention and oversight when using AI decision in specific sectors"*. Specifically, on civil liability, it considers that the rules must be adequate to ensure proper compensation for damages caused by AI-enabled devices and that mandatory insurance provisions may be necessary in this context.

[165] Also, when used for safety-critical applications.

b) For criminal liability, the HLG does no more than draw attention to the need for someone to be held liable for AI-caused damages, not elaborating on who should be responsible.

c) For consumer protection, it calls upon the EC to investigate whether current provisions are enough to protect consumers from being taken advantaged, explored and misled by AI *"and whether a mandatory consumer protection impact assessment is necessary or desirable"*.

d) On data protection the suggestions can be summed up into whether or not the GDPR offers an adequate level of protection. On this issue, the HLG draws attention to the fact that people may still be significantly affected by decisions that do not use personal data and there is currently no legislation offering the same level of protection that the GDPR offers for automated decision-making based on personal data (see our Chapter about the GDPR, on Part II) for equivalent automated decision-making that is not based on personal data.

e) Regarding legislation protecting against discrimination, (lack of) explanation requirements for AI in anti-discrimination laws are addressed, as well as the adequacy of enforcement mechanisms.

f) Cybersecurity merits no more than a general call for a complete adequacy assessment on current European Cybersecurity legislation.

g) In what concerns competition law, the HLG requests the Commission to start taking into consideration data held by companies in its *"assessment of market power for the purposes of applying rules on anti-competitive behaviour, abuse of dominance or (algorithmic) collusion, and when evaluating mergers"*.

The idea of a mechanism through which people may request the complete erasure of any public or private storage of data related to them as children is explored by the HLG. We do not know how the interplay between this mechanism and the GDPR would work. Furthermore, it is written in such a broad manner that it appears to call for a complete erasure of all data related to the person as a child. However, certain data must be stored for legal purposes (technically the data contained within the birth certificate is data collected from a person while he/she is still a child, so it should be erased under the new mechanism if written too broadly[166]).

[166] In fact, an for arguments sake, taking this provision at its most extreme, if in the exact same day a person reached legal majority he/she requested the erasure of all data in existence about him/her, he/she could "disappear" for all intents and purposes. The only piece of relevant data not affected by this general deletion

The HLG considers that the current policy cycle used by the European Institutions may not be adequate for AI. As such, it recommends systematic monitoring, period evaluation of the impact of regulatory measures and more and improved public consultations. Regulatory sandboxes (a common suggestion by both European Institutions and Member States as we have seen above and will further see below) are pointed as a solution for agile and better policy-making and developing methods to audit algorithms is identified as key to ensure that public authorities can identify and sanction unlawful use of AI. The need for redress mechanisms and ability to require human intervention are also considered.

Remarkably, the HLG appears to vehemently reject the notion of any new type of legal personality for AI systems or robots, stating that it is *"fundamentally inconsistent with the principle of human agency, accountability and responsibility, and to pose a significant moral hazard"*.

As expected, it calls for strong harmonisation with the absence of cumulative regulatory interventions at the national and with consistent and coherent enforcement throughout the EU. Somewhat in a contradictory manner, while rejecting cumulative national legislation because priority should be given to establishing the digital single market in one paragraph, in the paragraph immediately above the HLG states that *"harmonisation should however not preclude existing higher standards of protection at national level with regards to individuals, for instance regarding individuals' health and safety and consumer protection"*. We tend to agree with preposition one and disagree with preposition two. Minimum harmonisation tends to work quite badly in this type of situation as explained above. Thus, it is preferable to immediately enshrine a high level of protection in European legislation and then having it applied uniformly across Member States. The alternative could be a disorganised, hard to enter into and barely understandable set of legal rules, whose only similarity [and saving grace] is the fact that there is a minimum of protection that may not be infringed upon by Member States.

The HLG supports the implementation of guidance measures for market players working in AI solutions and certification and standardisation mechanisms. The Guidance should help market players understand how they can comply with the legal provisions in place for AI development and deployment and, we would add, should be updated along with any legislative measure adopted or relevant case-law by the ECJ. On certifications and

provision would be the request to have data deleted in itself, and even that one would then refer to someone who did not exist.

standards, the HLG considers that *"the development of co-regulation mechanisms for the certification of AI systems at EU level could counter fragmentation of standards, but could also help provide the means to assess the quality of an AI solution after deployment and possibly to decide which solution is best. Feedback received through the piloting phase of the Ethics Guidelines could help assess the necessity and shape of any such mechanism. As concerns standardisation, Europe should develop a clear strategy in terms of key standardisation fora and adequate resources. A European vision is needed on the main components of Trustworthy AI that may necessitate standards. Finally, it should be ensured that AI systems comply with mandatory sectoral certification requirements that already exist today"*. We would note that a co-regulatory approach would be in line with the HLG' position on a principle-based high-level approach to law-making regarding AI. The legislator could enshrine a few broad general principles and establish adequate tools to ensure that they were complied with (*e.g.* through the creation of Regulatory Authorities), but the specific methods for complying with would be decided by the relevant stakeholders. In fact, it is true that stakeholders have the most knowledge about the nuances of development in cutting-edge technologies such as AI and certainly have the incentive to work within the co-regulatory framework, as to avoid stronger legislative initiatives. However, this approach would have to be sector-specific and the level at its correct design and checks and balances applied would have to be very well though[167/168].

[167] *See*, Ira Rubinstein, "The Future of Self-Regulation is Co-Regulation", *in The Cambridge Handbook of Consumer Privacy, From Cambridge University Press*, Evan Selinger, Jules Polonetsky and Omer Tete (Cambridge: Cambridge University Press, 2018), 503-523

[168] We would like to note that EU is studying the potential set up of a 100 Billion fund to support the creation of European Companies that could compete with American and Chinese tech giants. The proposal for a European Future Fund, appears in a document prepared by European Commission officials for consideration by the new President of the EC, Ursula von der Leyen. The fund would invest directly in equity, a much more direct line of intervention than the one generally pursued by the European Union. According to POLITICO, who was able to obtain a copy of the internal document, *"the new fund would use money from the EU budget earmarked for venture capital, research funding and regional development. It would be overseen by the directorate general for research in discussion with five other departments, including those covering the budget and financial services"*. *See*, "Exclusive: Brussels eyes €100B wealth fund for 'European champions'", Bjarke Smith-Meyer, Lili Bayer And Jakob Hanke, accessed August 22, 2019, https://www.politico.eu/article/exclusive-european-commission-leaked-plans/; "European officials draft radical plan to take on Trump and U.S. tech companies", Bjarke Smith-Meyer et al., accessed August 22, 2019, https://www.politico.com/story/2019/08/22/europe-plan-trump-tech-companies-1472326.

Chapter IV – Development and AI-related initiatives in Non-EU Countries

§ 1. United States of America

The United States occupies the position of the world's technology leader. That status gives the country a sizable advantage in the AI race. Companies as Alphabet (Google), Amazon, Apple, Microsoft and Facebook possess a large user base and abundant stores of data from said users. In fact, even taking into account the natural fluctuations in the value of public traded companies, these five have dominated the list of most valuable companies in the world in the recent past. Microsoft is currently heading the list with a market cap of 1050 Billion dollars. Indeed, the US may count on its private sector more than any other region, having both the head-start, the knowledge and the means to develop AI without needing nearly as much government intervention as China, or even the EU. The values invested by tech giants in research and development frequently dwarf what is invested by most governments. In 2017, Amazon invested 16.1 Billion dollars, and Alphabet 13,9 Billion dollars. While not as impressive, the US federal government also invested more than 2 Billion dollars on AI research and development only on unclassified programs. That is, not including Pentagon and intelligence budgets, which tend to be areas where spending is quite high. Additionally, The Defense Advanced Research Projects Agency (DARPA) promised an additional 2 Billion dollars for AI research and development for a period of 5 years, running from 2017 onwards[169]. We should note that the US high level of expenditure on military matters may give the country an advantage also on AI development, especially if external contractors are hired to develop initially secret technologies but the know-how eventually makes it to commercial products. While, probably, not having the same high profile as the Chinese almost unified approach to commercial and military matters, it should not be downplayed as a factor that may bring successes to the US's hopes in AI development and deployment.

The US government created a Selected Committee on Artificial Intelligence to advise it on research and development priorities and to consider partnerships with

[169] *See*, "Defense Department pledges billions toward artificial intelligence research", Drew Harwell, accessed August 12, 2019, https://www.washingtonpost.com/technology/2018/09/07/defense-department-pledges-billions-toward-artificial-intelligence-research/.

academia and private stakeholders[170]. With this initiative the US intends to start laying foundations capable of ensuring; "*i) the maintenance of American leadership in AI; ii) support for the American workers; iii) the promotion of research and development and; iv) the removal of barriers to innovation.* The Selected Committee on Artificial Intelligence was also tasked with submitting a report *"to the President of the United States with recommendations on better enabling the use of cloud computing resources for federally funded AI Research and Development".* It shall also provide expertise to the American Technology Council. The two organisations updated in June 2019, the National Artificial Intelligence R&D Strategic Plan.

On 11 February 2019, the President of the United States issued Executive Order 13859: Maintaining American Leadership in Artificial Intelligence. This document, while not detailed in any sense of the word, or particularly dense or innovative has some interesting particularities. First, as the name implies, the US is the only country that claims, and arguably rightly so, leadership in the field of artificial intelligence. Therefore, while other countries are interested in achieving leadership in AI or in specific fields related to AI, the US is interested in maintaining the advantages it already enjoys. The document does contain some general provisions on the development of the workforce's skills and on ensuring public trust in AI and ensuring that its development leads to safe and secure uses of the technology. However, it is easy to understand that these objectives appear in the wider context of guaranteeing that there will be no trade barriers for American products (remarkably, and on a message clearly directed at China, while protecting American technologies from acquisition by strategic competitors and adversarial nations[171]). In fact, in 2018, the Foreign Investment Risk Review Modernization Act already called for stricter controls on foreign companies acquiring assets in American companies developing critical technologies. Eventually, the US may have to find a very delicate balance between defending a full free-market ideology and protecting key sectors and national security. If the notion of national security is excessively broad, American companies risk seeing other countries restricting their investment as a response to what may seem like restrictive

[170] In 2016 the US Government had already released three relevant reports on AI , the Artificial Intelligence, Automation, and the Economy; Preparing for the Future of Artificial Intelligence; and The National Artificial Intelligence Research and Development Strategic Plan (the last was updated in 2019).

[171] Besides their known conflict regarding American companies' IP, China also seems to be willing to block US-related actors from acquiring strategic assets, even if they are not Chinese. The manner in which the country hindered (while not outright blocking it, the fact is that it was made impossible) the acquisition of Dutch company NXP by Qualcomm is an example of this, even though both the European Union and the United States were okay with the operation. *See*, "China Blocks Qualcomm's Attempt to Buy a Dutch Chipmaker", Klint Finley, accessed August 1, 2019, https://www.wired.com/story/china-blocks-qualcomms-attempt-to-buy-a-dutch-chipmaker/.

measures by the Americans, if the opposite is true, the Country may be vulnerable to cyber-attacks to its most critical infrastructure. The 2016 elections may have already revealed some fragilities in the US's infrastructure, and it is certainly not keen on a repeat of those problems[172].

Regarding the issue of Trustworthy AI, one should note that, even if it does not hold centre stage as it does in the EU, there are indeed some measures and some attention being given to the issue. The 2019 update of The National Artificial Intelligence R&D Strategic Plan reserves a chapter for the issue of Ethical, Legal, and Societal Implications of AI, albeit severely lacking in depth. DARPA is working on XAI and it aims to *"create a suite of ML techniques that produce more explainable AI systems while maintaining a high level of learning performance (prediction accuracy)"*. Meanwhile, the national science foundation is working with Amazon to provide fairer AI with *"the goal of contributing to trustworthy AI systems that are readily accepted and deployed to tackle grand challenges facing society. Specific topics of interest include, but are not limited to, transparency, explainability, accountability, potential adverse biases and effects, mitigation strategies, validation of fairness, and considerations of inclusivity"*[173].

One should note that if the EU decides to set forth detailed and demanding rules for the development and deployment of AI in the Union (which actually appears to be the leader's current preferred strategy), there is the risk of creating/escalating a commercial conflict between the EU and the United States. It seems likely that, taking into account the

[172] *See*, "Executive Order 13859 of February 11, 2019: Maintaining American Leadership in Artificial Intelligence", Donald J. Trump, accessed June 10, 2019, https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/.

[173] *See,* "The National Artificial Intelligence R&D Strategic Plan: 2019 update", Select Committee on Artificial Intelligence and National Science & Technology Council", accessed August 15, 2019, https://www.whitehouse.gov/wp-content/uploads/2019/06/National-AI-Research-and-Development-Strategic-Plan-2019-Update-June-2019.pdf; "Artificial Intelligence for the American People: AI with American Values", United States Government, accessed August 15, 2019, https://www.whitehouse.gov/ai/ai-american-values/; "Explainable Artificial Intelligence (XAI)", DARPA, accessed August 15, 2019, https://www.darpa.mil/program/explainable-artificial-intelligence; "NSF Program on Fairness in Artificial Intelligence (AI) in Collaboration with Amazon (FAI)", National Science Foundation, accessed August 15, 2019, https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505651.

current administration's "America First" doctrine, that it will move to protect American companies' interests[174/175].

[174] The US answered France's attempt to tax Tech Giant's profits in the Country by immediately opening an investigation against the measure, which the US thinks may be discriminatory or unreasonable and burden or restrict US commerce. The investigation may result in sanctions against France. A plan to implement a similar set of measures in the UK is also, at the time of writing, creating issues in the relationship between the US and the UK. *See*, "US launches inquiry into French plan to tax tech giants", BBC News, accessed August 10, 2019, https://www.bbc.com/news/world-europe-48945828; "US to probe proposed French tech tax, concerned it 'unfairly targets American companies'", CNBC, accessed August 10, 2019, https://www.cnbc.com/2019/07/11/us-to-probe-proposed-french-tech-tax.html; "UK braves US ire by pressing ahead with tax on tech companies", Chris Giles and Claer Barret, accessed August 10, 2019, https://www.ft.com/content/41069548-a3d8-11e9-974c-ad1c6ab5efd1; "U.K. Presses Ahead With Tech Tax Plans Despite Rising Tensions With Washington", Sam Shead, accessed August 10, 2019, https://www.forbes.com/sites/samshead/2019/07/12/uk-presses-ahead-with-tech-tax-plans-despite-rising-tensions-with-washington/#100ad6116904.

[175] Even though the current administration's relationship with Tech Giants is not exactly a stable one. In fact, Donald Trump has previously called for a boycott against Apple products when the Company did not assist authorities in breaking the iPhone's encryption after a mass shooting. Currently, the relationship with the company appears good enough for the President to seek advice from Tim Cook on punitive tariffs against China (though last year, he suggest that Apple just move its manufacturing to the US if it wishes to avoid tariffs and it is not clear and, in fact, seems even unlikely that the President will heed Cook's advice). Trump has previously criticised Google, Facebook and Twitter for alleged liberal bias and Amazon for avoiding taxes. *See*, "Trump tweets support for far-right figures banned by Facebook", Brian Stelter and Oliver Darcy, accessed August 12, 2019, https://edition.cnn.com/2019/05/04/tech/trump-social-media-twitter-facebook/index.html; "Trump had a 'very good meeting' with Apple's Tim Cook about tariffs on iPhones, devices", Dalvin Brown, accessed August 20, 2019, https://eu.usatoday.com/story/tech/2019/08/19/trump-says-apple-ceo-tim-cook-makes-compelling-case-against-tariffs/2049030001/; "Donald Trump: Apple should make products in the US to avoid tariffs", The Guardian Staff, accessed August 12, 2019, https://www.theguardian.com/us-news/2018/sep/09/donald-trump-apple-should-make-products-in-the-us-to-avoid-tariffs; "Trump calls for Apple boycott", Jeremy Diamond, accessed August 12, 2019, https://edition.cnn.com/2016/02/19/politics/donald-trump-apple-boycott/index.html; "Donald Trump hits out at Facebook's Libra and bitcoin" Hannah Murphy, accessed August 12, 2019, https://www.ft.com/content/57692326-a452-11e9-974c-ad1c6ab5efd1; "Trump says he's 'watching Google very closely' after meeting with CEO", Colin Lecher, accessed August 12, 2019, https://www.theverge.com/2019/8/6/20756734/trump-google-anti-conservative-bias-claims-tweets; "Trump claims Google is suppressing positive news about him and 'will be addressed", James Vincent, August 12, 2019, https://www.theverge.com/2018/8/28/17790164/president-trump-google-left-wing-bias-claims; "Trump: Facebook, Twitter, Google are 'treading on very, very troubled territory and they have to be careful", Ryan Browne, accessed August 12, 2019, https://www.cnbc.com/2018/08/28/trump-accuses-google-of-rigging-search-results-in-favor-of-bad-coverage.html; "Trump criticizes Twitter in tweet, urges 'fairer' social media", Makini Brice and Susan Heavey, accessed August 12, 2019, https://www.reuters.com/article/us-usa-trump-twitter/trump-criticizes-twitter-in-tweet-urges-fairer-social-media-idUSKCN1RZ171; "Trump Attacks Amazon, Saying It Does Not Pay Enough Taxes", Michael D. Shear et al., accessed August 12, 2019, https://www.nytimes.com/2018/03/29/us/politics/trump-amazon-taxes.html; "Fact check: Trump falsely claims Google 'manipulated' millions of 2016 votes", Daniel Dale, accessed August 20, 2019, https://edition.cnn.com/2019/08/19/politics/trump-google-manipulated-votes-claim/index.html; "A Method for Detecting Bias in Search Rankings, with Evidence of Systematic Bias Related to the 2016 Presidential Election", Robert Epstein et al., accessed August 20, 2019, https://aibrt.org/downloads/EPSTEIN_et_al_2017-SUMMARY-A_Method_for_Detecting_Bias_in_Search_Rankings-EMBARGOED_until_March_14_2017.pdf.

§ 2. China

The Chinese philosophy and strategy regarding technological development frequently differs significantly from the one pursued by western countries in general and by the European Union in particular. One particular example would be the different approach to personal data protection. While the European Union stresses the right of its citizens to privacy and data protection including, but not limited to, privacy in the face of States' intrusion into people's private lives, China opts for a different approach. China and Chinese companies collect personal data on a massive scale, and, in fact, the country even uses AI to enable its government to keep peace and social harmony[176]. Of course, this approach, along with its large number of citizens gives China a significant advantage in collecting a large pool of personal data that can be key in the development of AI[177]. Though it would be wrong to think that China's only "ace in the hole" in the race for AI development is its large pool of data. Currently, China is a technological superpower in its own right, with the capability to develop world-class and highly refined technological solutions.

One measure that is particularly unpalatable for western tastes is the possibility of deploying a SCS. Frequently touted as an Orwellian-class surveillance nightmare fed by AI-enabled advanced surveillance[178], such a description is (presently) not accurate. Indeed, in 2014 a road map was published by China's State Council to establish a comprehensive credit score system by 2020 whose *"inherent requirements are establishing the idea of a sincerity culture, and promoting honesty and traditional virtues, it uses encouragement for trustworthiness and constraints against untrustworthiness as incentive mechanisms, and its objective is raising the sincerity consciousness and credit levels of the entire society"*[179]. However, currently in 2019 it is not possible

[176] It is a rerun of the eternal conflict between the right to privacy and keeping secrets and argument that "the one who is not guilty of any wrongdoing has not reason to be afraid".

[177] Though China allows private companies to massively collect data from its citizens and the government is relatively open with practice (doing the same itself), the fact is that access to public data in China lags behind other countries. China ranks 93rd globally for the openness of government data.

[178] Black Mirror's Season 03, episode 01, "Nosedive" presents an alternative reality where every social interaction is ranked from 1 to 5. Depending on a person's ranking it is possible to gain access to certain amenities such as cheaper housing or be in a special waiting list for flights. On the other hand, low rankings may restrict access to certain products and services (like only being able to rent low-end vehicles) or even job opportunities and certain medical treatments. The episode centres on Lacie Pound, a young woman trying to achieve a 4.5 ranking to access special conditions in renting luxury housing. Lacie is hit by a series of unfortunate circumstances, causing her ranking to plummet and ends up arrested but freer after she is stripped of her rating equipment.

[179] *See*, "China's Social Credit System: An Evolving Practice of Control", Rogier Creemers, accessed July 15, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3175792

to predict with confidence how a future central SCS will be implemented, what consequences for low and high credit citizens, what metrics will be used in the evaluation and how surveillance will be conducted. There are systems in place that act as probable precursors for the SCS such as the Joint Punishment System, designed to universally sanction people for rule-breaking in a determined area of their lives. The most wide-ranging implementation of this system is in sanctioning "*untrustworthy persons subject to enforcement*" (*i.e.* individuals who do not comply with a court's decision). Individuals may be subject to sanctions that may go from limitations to their ability to start a business, or to receive stock options to restrictions on their consuming habits, such as being unable to buy first-class tickets in trains or airplanes, visit luxury restaurants or resorts. In 2016, *"45 Party bodies, government departments and judicial institutions (…) concluded a memorandum of understanding to further clarify their respective roles in what had now become known as the Joint Punishment System"*[180]. The system was subsequently applied to other types of infraction, such as rail and air travel related.

Some Chinese cities also have pilot SCS systems in place. The city of Rongcheng assigns ratings from "A+++" to "D" to its citizens. High ranking citizens may get free medical check-ups, lower interest on loans and public recognition. Meanwhile, low ranking citizens may have their job prospects restricted. Rongcheng's "pilot" appears to be mainly rule-following based, as to lose points you would need to infringe existing laws and regulation or breach a contract. The city of Jinan introduced a SCS for dog owners, where they may find their pet "confiscated" and have to pass a test regarding the relevant regulations for retaining and raising dogs.

Private companies are also working on SCS (though their intention appears to use a supercharged traditional credit score, including social elements, to assess an individual's credit trustworthiness). One notable example is Ant Financial's (part of the Alibaba Group) Sesame Credit program which calculates scores based on five metrics *"(1) credit history, or records of past credit repayments; (2) behavioural trends, referring to someone's conduct when making purchases, processing payment, settling accounts and managing their finances; (3) ability to honour agreements, meaning having stable economic revenue and personal assets; (4) personal information, referring to the amount of verifiable and reliable information about themselves a member provides, and (5) social relationships, referring the extent to which one interacts with good friends and behaves in a friendly manner*

[180] *See*, "China's Social Credit System: An Evolving Practice of Control", Rogier Creemers, accessed July 15, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3175792

on the platform"[181]. About 387 Chinese cities accept this privately created SCS and offer perks to citizens with high credit. However, the scope and extension of the collaboration between Ant Financial and the Chinese government and how much data is provided by the former to the latter is unknown and debated. What is clear is that both the government and private companies may collect data without the "hindrance" from restrictive data protection laws such as the GDPR.

Furthermore, the differences in the fundamental rights and data protection frameworks mean that algorithms used to produce this type of assessments will not have to be subjected to the same degree of explicability requirements or accountability / auditability against bias and discrimination. This may prove to be an advantage in the development of AI due to lower compliance costs and absence of chilling effect. Nevertheless, the opposite could be argued, as the lower accountability requirements may fail to detect "bad data" and low-quality data may give rise to low-quality models, making the solution ineffective[182].

Still, and as stated above China does not depend only on data for its lofty AI ambitions. Development of AI and other emerging technologies is a key aspect of China's manufacturing strategy to be realised by 2025[183/184] and the country wants to be not a, but the world leader in the industry by 2030[185]. In 2016, the State Council's National Plan for Scientific and Technological Innovation During the Period of the Thirteenth Five-year

---

[181] *See*, "China's Social Credit System: An Evolving Practice of Control", Rogier Creemers, accessed July 15, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3175792

[182] *See*, "China's Social Credit Score – rating a people", Joanna Klabisch, accessed July 15, 2019, https://crossasia-repository.ub.uni-heidelberg.de/4177/1/2018_Juli_Social-Credit.pdf; "How Socioeconomic and Perceived Behavioral Patterns Impact Personal Zhima Credit Score in China's Credit System (DRAFT VERSION)", Jianyin Roachell, accessed July 15, 2019, https://www.researchgate.net/publication/327867876_How_Socioeconomic_and_Perceived_Behavioral_Patterns_Impact_Personal_Zhima_Credit_Score_in_China's_Credit_System_DRAFT_VERSION; "How The West Got China's Social Credit System Wrong", Louise Matsakislouise Matsakis, accessed July 15, 2019, https://www.wired.com/story/china-social-credit-score-system/; "China has started ranking citizens with a creepy 'social credit' system — here's what you can do wrong, and the embarrassing, demeaning ways they can punish you", Alexandra Ma, accessed July 15, 2019, https://www.businessinsider.com/china-social-credit-system-punishments-and-rewards-explained-2018-4; "China's Orwellian Social Credit Score Isn't Real", Jamie Horsley, accessed July 16, 2019, https://foreignpolicy.com/2018/11/16/chinas-orwellian-social-credit-score-isnt-real/; Samantha Hoffman, "Managing the State: Social Credit, Surveillance and the CCP's Plan for China", in *AI, China, Russia, and the Global Order: Technological, Political, Global, and Creative Perspectives*, in Nicholas D. Wright ed (Virginia: United States Department of Defense, 2018),42-47; Shazeda Ahmed, "Credit Cities and the Limits of the Social Credit System", in Nicholas D. Wright (ed.), *AI, China, Russia, and the Global Order:..*,48-54

[183] In fact, China's 2006 National Medium- and Long-Term Plan for the Development of Science and Technology already established plans for the development of technologies such as smart robots or virtual reality.

[184] In accordance with the Made in China 2025 strategic plan (2015).

[185] In accordance with the China's New Generation AI Development Plan (2017).

Plan established AI as the main priority in the next generation of developing information technologies. The areas of "*natural human-computer interaction, especially intelligent perception and cognition, virtual-physical integration and natural interaction, and semantic understanding and smart decision-making*" were selected as key for the Chinese strategy requiring takes measures aimed at *"vigorously developing big data-driven human-like intelligence technologies and methods; making breakthroughs in human-centric human-machine fusion theories, methods and key technologies and developing related equipment, tools and platforms; and making breakthroughs in human-like intelligence based on big data analysis and achieving humanlike vision, hearing, speech and thinking to support AI-driven industrial development and demonstrative applications in key sectors such as education, office and healthcare*[186]*".*

With a bubbling e-commerce ecosystem and a third of world's tech unicorns, China possesses a high level of digitalisation and (sometimes very close) government support (including through research) to back its private sector[187]. The Chinese internet market is one of the largest in the world, with 800 Million connected users who can serve as both consumers and producers of data for AI training.

To achieve its AI development and deployment objectives the country plans to: "*i) build an open, collaborative AI technology innovation system, cutting edge in terms of basic theory, key common technologies, innovation platforms, and high-end talent; ii) foster a high-end and efficient smart economy, develop the emerging AI industry, promote industrial upgrade and create AI innovation areas; iii) build a safe and supportive smart society, develop efficient and intelligent services including for social governance, enhance public safety and security based on AI and promote sharing and mutual trust in social interactions; iv) strengthen military and civilian integration in the field of AI, promote the two-way conversion of AI technology between military and civilian sectors and jointly build up and share military and civilian innovation resources; v) construct a ubiquitous, safe and efficient intelligent infrastructure system and strengthen the upgrading of infrastructures such as networks, big data and high-performance computing; vi) plan significant projects, and strengthen overall coordination.*

There are also regional projects that may be of central importance for the expansion of the AI ecosystem. Cities such as Beijing, Shanghai, Shenzhen and Zhongguancun (aka the Chinese Silicon Valley) are known technology hubs. Large market

---

[186] *See*, China Institute for Science and Technology Policy at Tsinghua University, C*hina AI: Development Report 2018* (Beijing; Tsinghua University, 2018): 70ff.

[187] For example, the Chinese government gave SenseTime access to its databases containing citizens' data. *See*, "Meet the World's Most Valuable AI Startup: China's SenseTime", Bernard Marr, accessed July 5, 2019, https://www.forbes.com/sites/bernardmarr/2019/06/17/meet-the-worlds-most-valuable-ai-startup-chinas-sensetime/#7bc932e7309f.

players like Alibaba, Tencent, SenseTime and Baidu can give American Tech companies a run for their money and dwarf the European technological sectors, arming China with a potential advantage on existent know-how and market power for AI development.

In May 2019, a group of prestigious universities, including the Beijing Academy of Artificial Intelligence, Peking University, Tsinghua University, Institute of Automation and Institute of Computing Technology in Chinese Academy of Sciences and influential Chinese companies including Tencent, Alibaba and Baidu released the "Beijing AI Principles". These 15 principles are divided in principles for Research and Development, Use and Governance. In these principles, we may find some familiar concepts, including the necessity to respect human privacy, dignity, freedom, autonomy, and rights or to consider potential ethical, legal, and social impacts and risks in AI R&D. The concept of trustworthiness (in this case a trustworthy system) is also called upon, and open platforms are promoted. The principles also contain the need for informed consent for AI use and the need for reasonable data and service revocation mechanisms (though no explanation). Nevertheless, they are quite vague, not endorsed by official authorities and we do know that some of the concepts will not necessarily be interpreted in the same manner as they are in the EU (such as human privacy, freedom or informed consent). However, some room appears to exist to engage with Chinese authorities and stakeholders to achieve some harmonised rules on AI development and deployment.

Like the European Union and the United States, China currently lacks qualified human resources to turn its AI ambitions into reality. The 2015 version of the Catalogue of Student Majors in Vocational and Technical Education for Regular Institutes of Higher Education stressed the importance of AI-related studies. By 2016, programs related to AI and robotics were available in 300 higher education vocational institutes. However, there are questions about whether the Chinese education system will be able to respond to the challenges in a sufficiently swift manner. Furthermore, the Chinese Communist Party will have to balance the need for automation in its industries with the impact in the job market, which can lead to social unrest, that it seeks to avoid. Even if its GPD can rise as much as 26% due to AI deployment, the rise in inequality may offset the economic gains. China being the world's nation with the highest level of automatable tasks, can be particularly affected by this problematic particularity. Ineffective allocation of resources and difficulties arising from the need for researchers to keep good relations with the government may also create and added difficulty for China to maintain talent in an industry where outside

competition will be fierce[188]. Chinese researchers are also excessively concentrated in specific areas and may not be able to offer a sufficiently wide-ranging cover to enable the creation of a complete leader in AI.

Additionally, the Chinese data ecosystem is not sharing-friendly, with a lack of unified standards and cross-platform sharing hindering it. The Country is also dependent on foreign software libraries for ML development such as TensorFlow (open-source, but initially developed by Google), PyTorch (also open-source but its main contributor is Facebook) or Keras (open-source, but originally developed by François Chollet a French Google Engineer). China is also unable to design and build high-value semiconductor chips

[188] Even so, if we go by quantity, in 2015, China and the United States produced nearly 10.000 scientific papers on AI-related fields, while Germany, the UK, India and Japan combined were only able to produce approximately half of the scientific output. However, when we look at citations outside of China, it still lags behind other countries, showing that its research may be less influential. *See*, Dominic Barton et al., "*Artificial Intelligence: Implications for China"* (New York: McKinsey Global Institute, 2017): 4ff.

(GPUs and specialised AI chips) with the same quality as the US, making it vulnerable to restrictions and limitations in importing and/or providing technology by the US[189/190].

[189] The Chinese mobile phone manufacturer ZTE was put in serious peril of bankruptcy when the U.S. Department of Commerce announced a seven-year ban on exports by US of vital components for the ZTE assembly line, due to the company's dealings with Iran and North Korea. The decision was later reversed after (a highly unusual) political intervention by US President Donald Trump following discussions with the Chinese President Xi Jinping. Another interesting and ongoing case was Huawei's a multinational technology company, specialized in telecommunications (even though it is known for its leading position in the mobile phone market) who also suffered bans that could have killed its international phone business by preventing its use of the Android OS (at least the non-open source parts, including Google Apps) and import of vital components (in line with ZTE). The Huawei case was handled in a more political fashion, with the authorization to let the Treasury Department put Huawei on list banning business from US companies coming from an executive order by the President of the United States while in ZTE's case the placement of the list of banned companies arose from an investigation by the United States Department of Commerce. As abovementioned the case is still ongoing, but the restrictions were eased after another intervention by the Chinese President, albeit as of this writing, Huawei is only, for example, able to use Android OS with Google Apps, on existing handsets and not on new devices. "China seeks semiconductor security in wake of ZTE ban", Edward White, accessed June 17, 2019, https://www.ft.com/content/a1a5f0fa-63f7-11e8-90c2-9563a0613e56; "US companies banned from selling components to ZTE", Brian Heater, accessed June 17, 2019, https://techcrunch.com/2018/04/16/u-s-companies-banned-from-selling-components-to-zte/; "Trump orders U-turn over sanctions against Chinese telecoms group", Sam Fleming and Shawn Donnan, accessed June 17, 2019, https://www.ft.com/content/8fc1b404-56c4-11e8-b8b2-d6ceb45fa9d0; "ZTE fined $1 billion", Danny Crichton, accessed June 17, 2019, https://techcrunch.com/2018/06/07/zte-fined-1 billion/?guccounter=1&guce_referrer_us=aHR0cHM6Ly93d3cuZ29vZ2xlLnB0Lw&guce_referrer_cs=kFeknkKSA2-T3R3ISgwUDA; "Huawei Can't Guarantee Mate 30 Will Ship With Android, But Will Be 'Ready' With Alternative OS", Ben Sin, accessed July 30, 2019, https://www.forbes.com/sites/bensin/2019/07/30/huawei-cant-guarantee-mate-30-will-ship-with-android-but-will-be-ready-with-alternative-os/#354066445d09; "Trump says he'll ease Huawei restrictions, but no one's sure how", Colin Lecher, accessed July 30, 2019, https://www.theverge.com/2019/7/3/20679998/trump-huawei-trade-announcement-restrictions; "Executive Order on Securing the Information and Communications Technology and Services Supply Chain", Donald J. Trump, accessed July 30, 2019, https://www.whitehouse.gov/presidential-actions/executive-order-securing-information-communications-technology-services-supply-chain/; "What does Huawei's trade ban mean for your Huawei or Honor phone?", Dan Grabham, accessed October 25, 2019, https://www.pocket-lint.com/phones/news/huawei/148102-what-does-huawei-s-google-ban-mean-for-your-huawei-or-honor-phone; "Huawei: Record 200 Million Devices Shipped As Android Future Resolved", Zak Doffman, accessed October 25, 2019, https://www.forbes.com/sites/zakdoffman/2019/10/24/huawei-record-200-million-devices-shipped-as-android-future-resolved/#34d2a2446f40.

[190] *See*, Max Craglia (Ed.), *Artificial Intelligence: A European Perspective* (Luxembourg: Publications Office of the European Union, 2018): 45-51; "Scenarios and Potentials of AI's Commercial Application in China", Deloitte, accessed June 20, 2019, https://www2.deloitte.com/content/dam/Deloitte/cn/Documents/innovation/deloitte-cn-innovation-ai-whitepaper-en-190118.pdf; Jeffrey Ding, "The Interests Behind China's AI Dream", in Nicholas D. Wright (ed.), *AI, China, Russia, and the Global Order*:..,37-41; Benjamin Angel Chang,"AI and US-China Relations", in Nicholas D. Wright (ed.), *AI, China, Russia, and the Global Order*:.., 107-111; China Institute for Science and Technology Policy at Tsinghua University, C*hina AI: Development …,* 62ff.; Dominic Barton et al., "*Artificial Intelligence: Implications for China"* (New York: McKinsey Global Institute, 2017): 4ff., "Chinese Interests Take a Big Seat at the AI Governance Table", Jeffrey Ding, Paul Triolo, and Samm Sacks, accessed June 10, 2019, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/chinese-interests-take-big-seat-ai-governance-table/; "AI Policy – China", Future of Life Institute, accessed June 29, 2019, https://futureoflife.org/ai-policy-china/; "Beijing AI Principles", Tsinghua University et al., , accessed June 29, 2019, https://www.baai.ac.cn/blog/beijing-ai-principles; Yujia He, *How China is preparing for an AI-powered Future* (Washington D.C.: Woodrow Wilson International Center for Scholars, 2017); 1ff., Sophie-Charlotte Fischer, *"Artificial Intelligence: China's High-Tech Ambitions"* Center for Security Studies Analysis in Security Policy 220 (2018): 1-4; "Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security", Gregory C. Allen, accessed July 3, 2019,

§ 3. Japan

Currently the third largest economy in the world, with notable companies such as Sony, Nintendo, Toshiba or the NEC Corporation, Japan intends position itself as a relevant player in the global AI race. The country was quick to identify the importance of AI development for its future and, by 2017, already had an AI strategy. Through the development of IoT, Big Data, AI and robotics Japan aims to grow its GDPR by about 10% and create 272.7 Billion dollars in added value by 2020.

Japan's Artificial Intelligence Technology Strategy is not very long or overly detailed, but it contains the main challenges and issues one would expect in a document of this nature: *i)* fostering R&D in AI-related fields in an attempt to catch up to the United States and China; *ii)* training more professionals specialized in AI to answer the market's needs; *iii)* making data available for development while taking into account such *"as reliability, security, system flexibility, personal information protection, balance between oligopoly and utilization and application of data, and coordination among data"*; *iv)* investing in new technologies such as 5G and specialized chips necessary for the development and deployment of AI; *v)* guaranteeing funding for companies and start-ups.

In its strategy, Japan identified four key areas for AI development, namely productivity, health, medical care, and welfare, mobility, and information security.

On productivity, Japan wants to universally integrate AI solutions in the manufacturing, distribution and services industries. With the aid of AI, it also aims to foster creativity (maybe by relieving people of repetitive tasks), *"leading to a society where innovative services and products can be continuously created"*.

Regarding medical care and welfare this is a particularly delicate issue for Japan as the country suffers from chronically low birth rates and an ageing population. The Japanese AI Strategy itself draws attention to the fact that by 2030, over 40% of the

https://s3.amazonaws.com/files.cnas.org/documents/CNAS-Understanding-Chinas-AI-Strategy-Gregory-C.-Allen-FINAL-2.15.19.pdf?mtime=20190215104041; "What People Get Wrong About China and Artificial Intelligence", Jonathan Vanian, accessed July 12, 2019, https://fortune.com/2019/07/09/china-data-artificial-intelligence/; "The Chinese AI innovation chasm", Michael R. Wade and Amanda Bris, accessed July 12, 2019, https://www.imd.org/research-knowledge/articles/the-chinese-AI-innovation-chasm/; "What you may not understand about China's AI scene", Karen Hao, accessed June 13, 2019, https://www.technologyreview.com/f/613296/what-you-may-not-understand-about-chinas-ai-scene/; "Why does Beijing suddenly care about AI ethics?", Will Knight, June 13, 2019, https://www.technologyreview.com/s/613610/why-does-china-suddenly-care-about-ai-ethics-and-privacy/; "Cold water hits China's AI industry", Louise Lucas, accessed July 10, 2019, https://www.ft.com/content/973bfc08-a15f-11e9-a282-2df48f366f7d.

Japanese population will be elderly and that Japan is the most rapidly ageing society in the world.

Mobility is also an obvious key area for Japan, as the country possesses a reasonably sizable and advanced automobile industry. General considerations regarding safe, environmentally friendly travel and freeing up time to pursue other activities while travelling makes up the core of the mobility section of Japan's AI strategy. Still, and very interestingly, Japan includes virtual travelling on this section, even if it is not delved into too deeply.

Concerning information security, Japan considers that *"not only will reliability and stability be emphasized, but the confidentiality of technology will also be emphasized, and technological development will progress"*[191].

In its favour, we have the fact that, Japan is only behind the United States and China in number of AI-related patents filled. In fact, of the 20 companies with most patent filled in the area, 12 are Japanese[192]. In its 2018 budget, Japan marked 357.3 million dollars for *"developing robot related technologies and AI chips for next generation computers"* and 178.2 million dollars for *"AI in medical data management and pharmaceutical research"*[193].

However, some consider that this may not be enough and that, besides the issue of qualified professionals, Japan may also be currently lacking adequate funding (Japan's budget for AI pales in comparison to the US' or China's). Furthermore, even if Japan possesses large amounts of data it is not always in usable form and, similar to the EU's, its data protection rules are stricter than in the US or China[194]. In addition, in the software

191 *See*, "Artificial Intelligence Technology Strategy", Strategic Council for AI Technology, accessed August 13, 2019, https://www.nedo.go.jp/content/100865202.pdf.

192 *See*, World Intellectual Property Organization, *WIPO Technology Trends 2019: Artificial Intelligence* (Geneva: WIPO, 2019: p. 32ff.

193 *See*, "AI In Japan Opportunities For Swiss Companies", Swiss Business Hub Japan, accessed August 14, 2019, https://www.s-ge.com/sites/default/files/publication/free-form/s-ge-artificial-intelligence-report-japan.pdf.

194 We should note that the EU and Japan agreed to recognize their level of protection regarding data protection as equivalent, opening up the free-flow of data between both blocks, a fact that can contribute to lighten the burden of stricter data protection rules by making available larger quantities of data that can be, potentially, quite diverse in nature. See, "Commission Implementing Decision 2019/419/EU of 23 January 2019 pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council on the adequate protection of personal data by Japan under the Act on the Protection of Personal Information", European Commission, accessed February 12, 2019, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019D0419&from=EN; "European Commission adopts adequacy decision on Japan, creating the world's largest area of safe data flows", European Commission, accessed February 12, 2019, https://europa.eu/rapid/press-release_IP-19-421_en.htm.

realm, the Country does not enjoy the same as accolades as it does in hardware, which may limit its AI capabilities[195].

Chapter V – Development and AI-related initiatives in EU Member States

§ 1. France [196]

The French economy is the third largest in the European Union (behind Germany and the UK)[197]and the country is held in the highest regard in the European diplomatic circles. Furthermore, France is an extremely influential actor in policy and law-making in the European Union and has been working to foster European initiatives in the Digital Single Market in general and in AI specifically.

The country has been a pioneer in the development and strategy-building regarding AI in Europe. France has had regulation on automatic decision making since the 1970s and, in fact, you can see the influence of French legislation, like the Digital Republic Act of 2016, in the solutions adopted at a European level such as the GDPR[198]. It is quite interesting to note that the Villani Report (in which the French AI strategy "AI for Humanity" is largely based) focuses heavily on European cooperation[199]. In fact, when

---

[195] *See*, "Japan falling behind in artificial intelligence, warns SoftBank founder", Kana Inagaki, accessed August 13, 2019, https://www.ft.com/content/cab0936c-a940-11e9-984c-fac8325aaa04; Toyoaki Nishida, *"The Best of AI in Japan — Prologue"*, AI Magazine, 33(2): 108-111; "Japan aims to produce 250,000 AI experts a year", Minako Yamashita, accessed August 12, 2019, https://asia.nikkei.com/Economy/Japan-aims-to-produce-250-000-AI-experts-a-year; "AI Policy – Japan", Future of Life Institute, accessed August 13, 2019, https://futureoflife.org/ai-policy-japan/.

[196]. In our analysis of the development and AI-related initiatives in EU Member States we will make continuous use of the following sources: "Government AI Readiness Index 2017", Oxford Insights, accessed July 5, 2019, https://www.oxfordinsights.com/government-ai-readiness-index; "ITUC Global Rights Index from the 2018", International Trade Union Confederation, accessed July 5, 2019, https://www.ituc-csi.org/IMG/pdf/ituc-global-rights-index-2018-en-final-2.pdf; "WJP Rule of Law Index 2019", World Justice Project, https://worldjusticeproject.org/sites/default/files/documents/WJP-ROLI-2019-Single%20Page%20View-Reduced.pdf; "Special Eurobarometer 447 (regarding users concerned about data collected about them in the internet) European Commission, accessed July 5, 2019, http://ec.europa.eu/information_society/newsroom/image/document/2016-24/ebs_447_en_16136.pdf.

[197] "GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

[198] *See,* "State-of-the-Art Report | Algorithmic decision-making", algo:aware, accessed July 5, 2019, https://www.algoaware.eu/wp-content/uploads/2018/08/AlgoAware-State-of-the-Art-Report.pdf

[199] In fact, a great number of highly relevant reports and position papers were produced by French initiative as France Stratégie's Intelligence artificielle et travail (which contributed to the Villani Report), CNIL's "How Can Humans Keep the Upper Hand? the ethical matters raised by algorithms and artificial intelligence" or FranceIA's France intelligence Artificielle.

unveiling the Country's AI strategy French President Emmanuel Macron stressed the importance of European cooperation.

Regarding private sector actors working on AI, the number of start-ups increased 14-fold since 2000 and company contributions to the development of the technology are increasing at a very high rate[200]. In addition, Google and Facebook already have AI researchers in the country and IBM, Samsung and Fujitsu will soon follow suit[201]. Airbus, even though its headquarters are in the Netherlands, has its main office and some their key production facilities in France. French companies like BNP Paribas and the AXA group are also investing on introducing AI to their traditional markets. Reaching and conduction interviews with this rich ecosystem of stakeholders will be add considerable value to our report.

§ 1.1. The Villani Report (For a Meaningful Artificial Intelligence: Towards a French and European Strategy[202]

Due to its deepness, quality and complexity the Villani Report must be analysed separate from the general considerations about AI implementation in France. The first thing that immediately should be taken into account is that the Villani Report intends to build a French and European AI strategy, not just a French one. Immediately, it is clear that France is looking outwards, to the European Union to create a coordinated strategy and implement a coordinated approach into AI development. Cooperation between Member States is, in fact, frequently stressed, particularly between France and Germany (with the report also showing some interest in the work on robotics done in the North of Italy). A study of the possibilities of harmonization of AI policies in key sectors is recommended and the investment in European AI infrastructure is also pointed as a priority.

The general worries about creating, attracting and keeping talent in AI-related fields are also present in the Villani Report. Specially the Report considers that public researchers

[200] *See*, "How France became the place to be for AI startups", TechStartups Team, accessed May 7, 2019, https://techstartups.com/2019/01/16/france-became-place-ai-startups/
[201] *See*, "France wants to become an artificial intelligence hub: attracting talent, fostering public research and AI startups", Romain Dillet, accessed May 7, 2019, https://techcrunch.com/2018/03/29/france-wants-to-become-an-artificial-intelligence-hub/
[202] *See*, "For a Meaningful Artificial Intelligence: Towards a French and European Strategy", Cédric Villani et al., accessed May 7, 2019, https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf.

will need better conditions to avoid losing them to private companies, namely in terms of freedom to pursue projects and financial compensation. Fostering the training of new professionals (tripling the number of AI graduates) and synergies between business and academia are also priorities that must be addressed.

Labour and the labour market are also highlighted in the Villani Report. The report appears to reject the darker predictions about AI's impact in the labour market, but also does not sign below the ones that state that seamless transition is the most likely scenario. It argues that moments of economic transition tend to bring with them difficulties and the best approach is to face the issue head-on. Adapting the workforce to tasks that will not be automated in the future. Skills that allow humans to work in a complementary in manner with AI are a must. Still, humans must remain in control and not be diminished by AI. To guarantee this the Report proposes to work with employees and employers alike. If needed, legislative reforms can be implemented that ensure working conditions and the good type of complementary work between AI and humans, a possible change in rules about collective bargaining is also addressed. Furthermore, the report proposes a public think-thank for transforming work. Inclusion is also a relevant factor. Women are grossly underrepresented in ICT in general, and AI specifically. The report proposes measures to help in countering this issue, especially in the classroom and workplace. On diversity in general, it also proposes funding to develop including and non-discriminatory AI and avoiding that some people are "left-behind" by the digitation and AI technologies, including through public services and promotion of AI-based social innovation.

Notably, the environment is the subject of an entire chapter of the Villani report. Appropriately, the Report draws attention to the fact that AI (and other emerging technologies such as bitcoin) may cause the world's energy needs to skyrocket. On the other hand, AI may also give valuable contributions to greener and more effective energy. Thereby, releasing ecological data for AI-development in this area is relevant (below), supporting greener AI-technologies and putting the issue on the international agenda (with the backing of the European Union) are also key.

On data, the Villani Report draws attention to the imbalance between the Americans Google, Amazon, Facebook, Apple and Microsoft, and the Chinese Baidu, Alibaba, Tencent and Xiaomi and every other player in the AI-development market. This fact according to the report, creates a data imbalance that puts these companies in great advantage. Since governments in the European Union do not implement some "stronger"

and liberty-restricting policies as China or Russia they are not being able to keep data in-house and, in fact, most of the Europeans' data is collected by American companies. A suggestion that will also appear in the German AI Strategy, the Villani Report suggests an aggressive policy at promoting data access. It calls upon public authorities to introduce new ways to make data available to relevant stakeholders, including encouraging economic players to share and pool their data. The Villani Report also calls for prompt and proper implementation of the (Proposal for a revision of the) [new] PSI Directive[203/204] and for more flexibility on copyright to allow for text and data mining. The implementation of non-proprietary, interoperable standards and investing in safety and security in AI is also key. This to avoid a scenario where the European Union becomes a digital colony of the United States and China. The Villani Report regards that if the GDPR had existed 20 years ago, European companies would have been more successful in competing with American giants for market opportunities. The Report also proposes the creation of collective rights concerning data, to avoid discrimination of certain sections of the population, even when there is no personal data being processed. Combining the current DPIA with new Discrimination Impact Assessments (hereinafter "DIA"), that can compel programmers to think about mitigating measures against discrimination (in the case of DIAs even when there is no personal data) is also suggested.

The Report warns against spreading French efforts to too thin and suggests that the Country focus on 4 sectors, healthcare, environment, transport-mobility and defense-security[205]. These are key sectors, that can energize relevant stakeholders and that need a strong impetus from the State[206]. It considers that the first battle in the "AI war" was already fought and won by major platforms (focusing on personal data), but in the second, for sector-specific data Europe could still have a major role to play. Raising the profile of French AI players is very important to achieve success in this area.

---

[203] Regarding this particular legal act, the Report considers that *"opening-up of access to certain sets of data—on a case-by-case basis and depending on the different sectors—on grounds of general interest"*, should be studied and it could take the form of access by public administration that then would feed a public database or direct access by other private stakeholders. However, it calls for caution if implementing this type of measure, as a general regime of free access to private data seems undesirable.

[204] Realised through the Open Data and Public Sector Information Directive.

[205] Since it is a very specific French issue, it falls out of the scope of this Thesis and would make our analysis of the Villani Report and the France Strategy extremely detailed (which is not supposed to be the case) we will not go into the 4 specific sectors in depth.

[206] The Report considers that the State should lead by example on the issue of AI-development and implementation.

A peculiar idea in the Villani Report is the suggestion that exceptions to public procurement rules could be introduced to give priority to European companies when there is a great imbalance in market conditions.

According to this document our high standards and tradition of protecting individual freedoms should be considered as a strategy and differencing factor and not a hindrance. It believes that is possible to be an industry leader without renouncing them.

A plethora of ethical questions are considered. Most of them like bias, explaining AI's decisions and use of AI for war and policing have already or will be referred elsewhere in this Thesis. We would like to quickly point one important point and one very interesting suggestion though: the important point is that the concept of ethics by design appears in this report, in line with the European Commission's communications. The interesting suggestion is the creation of a group of European advisory bodies ethical development of AI and digital technology, (roughly) modeled on the EDPB (which replaced the WP29). If, in the future, there is AI-specific legislation that needs the intervention a supervisory authority working specifically on this matter, maybe said advisory bodies could (in fact, should) turn into full-fledged Regulators (approaching the EDPB, which is composed by supervisory authorities on data protection).

§ 2. Germany

Germany is the largest economy in the European Union (the Country is responsible for more than one fifth of the block's GDP[207]) and, similarly to France, is held in the highest regard in the European diplomatic circles. Germany is highly influential in policy and law-making in the European Union. As the most populous Member-State in the EU also possesses the largest number of MEPs. Moreover, the President-elect of the European Commission comes from the German government (was the German defence minister), Germany's policy preferences and its stakeholders' opinion will be a key force in modelling present the present and the future of the EU's AI landscape.

[207] *See*, "GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

Germany was one of the first countries to adopt and AI strategy[208]. Germany's AI Strategy is quite complex and detailed (even though not as deep as the Villani Report), containing high-level goals and low-level projects and measures to ensure those goals are met. The strategy frequently refers to previously implemented measures and explains how they will be expanded and adapted to meet the needs of AI development. Germany will rely heavily on its world-class research capabilities, advance its existing partnership (namely across the EU, where France should be highlighted) and work on fast-tracking the development to product while making sure that promising technologies are not lost before they reach the marketplace. The areas where Germany wants to invest are quite broad, going from agriculture to space. On that note, the country will promote cooperation between companies that are working on AI (within the limits of what is legally permissible, namely by competition law), promote regulatory sandboxes, test beds and test labs so that companies can develop their products easily but in a compliant manner. Start-ups and companies that are scaling-up are also addressed as a priority through the creation of tools to support their growth or plain financing as is promoting the use of AI in public administration.

Balancing the interests and rights of workers and companies in the context of AI development is a difficult challenge that the Germany AI strategy promises to meet head-on. Measures in this area include investing in *"technologies designed to be socially compatible and in giving workers the skills they need"*. Indeed, adapting the current educational framework to the needs of the future workforce is one of the key aspects of the strategy. A new Workers' Data Protection Act is also being studied *"(within the limits of the GDPR) to protect employee's data in the age of AI"*.

Germany will create a new AI observatory that will follow the developments of this technology around the world to assess the impact in the various relevant areas. The government is also working on manners that make auditing AI companies possible, the objective is to *"set benchmarks for employment, technical design, human-machine interfaces, health and safety, and data protection"*.

Europe has some of the most demanding standards in privacy and data protection standards across the world. Undeniably, there are some challenges arising from it, which

[208] *See*, "Artificial Intelligence Strategy", Germany's Federal Government, accessed June 10, 2019, https://www.ki-strategie-deutschland.de/home.html?file=files/downloads/Nationale_KI-Strategie_engl.pdf.

we shall analyse below. Suffice to say the companies in the EU do not have the same broad access that leading companies in China (arguably) have to citizen's data[209]. But in AI development data is king. Aware of this fact, Germany wants to make robust, high-quality data available to companies in sufficient quantity. Promoting the European data space, the flow of data from the public to the private sector and data-sharing agreements between companies are key priorities. Studying the impact of said data-sharing agreements in competition law is also pointed as a necessity. Standardization, interoperability of data and research in areas related to anonymization, and development of AI with the use of lower quantities or no personal data is also pointed as an area of research to give primacy to.

The country also wants to promote cooperation between data protection authorities and relevant stakeholders to build guidelines on the use of personal data in AI development. All of these measures should be implemented in a way that respects the current European framework on data protection. In the strategy it does not seem like there is any desire for a special and less strict regime for the processing of data for AI-development. In terms of regulatory development Germany is in line with the Commission's proposals, including on ethics by design[210].

The Country also plans to inject €3 billion in the economy for AI development and expects to secure further financing from private entities. Like France's the Germany AI strategy also bets on European cooperation as a differentiating factor that will help the Country in the race against third countries like the United States and China.

The EU's largest economy can depend on the power of their homegrown multi-national companies that are already investing in AI such as SAP, Bosch, Siemens, DHL and car manufacturers BMW, Mercedes-Benz and AUDI. Foreign companies like Google are also installing research centres dedicated to AI in Germany. Furthermore, German cities are having a great degree of success in fostering the creation and implementation of

---

[209] *See*, "China's edge in the tech race is vast amounts of data", Shafi Musaddique, accessed July 10, 2019, https://www.cnbc.com/2018/11/30/chinas-edge-in-the-tech-race-is-vast-amounts-of-data.html, "China's greatest natural resource may be its data", "Kai-Fu Lee", accessed July 16, 2019, https://www.cbsnews.com/news/60-minutes-ai-chinas-greatest-natural-resource-may-be-its-data-2019-07-14/. Even if Germany's business tissue is quite impressive, it does not have the same power in this area as the US's. In fact, last year, 36.2% of the AI-related patents filed in Germany came from American companies, while just 18.1% came from German companies. Japanese companies filled 13,3% of patents, South Korean ones 3,7%, French 3,3% and 2,9 from Dutch companies.

[210] *See*, "A Look at Germany's AI Strategy", Fabian Schmidt, accessed 15 July, 2019, https://iiot-world.com/artificial-intelligence/a-look-at-germanys-ai-strategy/; "AI Made in Germany — The German Strategy for Artificial Intelligence", C. Koch, accessed 15 July, 2019, https://towardsdatascience.com/ai-made-in-germany-the-german-strategy-for-artificial-intelligence-e86e552b39b6.

companies dedicated to AI research. Berlin alone is home to hundreds of those companies that generate hundreds of millions of Euros in revenue.

§ 3. United Kingdom [211]

At the current stage it is extremely difficult to predict what results will be achieved when the Brexit process is over. At the time of writing an extension until January 2019 had been agreed between the European Union and the UK. In addition, more extensions are not at all outside of the realm of possibility, if the current agreement between the EU and the UK does not find the support in the British Parliament (or even the EP) or if a General Election in the UK changes the arithmetic of British politics . Furthermore, the UK can always revoke Article 50 TEU, in accordance with the Court of Justice's case-law. Boris Johnson, a hard Brexiter, was recently appointed[212] as Prime Minister of the United Kingdom. Johnson promised to leave the EU on 31 of October deal or no deal (though is default position is deal) but failed to do so. Currently Johnson is trying to call a General Election, expecting that it would deliver him the majority he needs to force his Brexit deal through Parliament[213].

With the above political situation in mind, the UK is still a Member State of the European Union and enjoys all its rights as one until the moment it leaves the bloc. The UK is also, currently, the second largest economy in the European Union and a large technology and AI development Hub. The UK tops Government AI Readiness Index from Oxford Insights, which analyses the current capacity of OECD governments to absorb - and exploit - the innovative potential of AI.

British companies and start-ups have been consolidating their position on the world's AI market. British founded Deepmind was bought by Google for €500 million in

---

[211] *See*, GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

[212] Boris Johnson was chosen by the Members of the Conservative Party (who currently holds government) to replace Theresa May.

[213] *See,* "Irish minister says Boris Johnson's Brexit stance 'quite alarming'", Lisa O'Carrol, accessed June 13, 2019, https://www.theguardian.com/politics/2019/jul/25/irish-minister-says-boris-johnsons-brexit-stance-quite-alarming; "Boris Johnson vows to ditch backstop and scale up no-deal plans", Peter Walker, accessed June 13, 2019, https://www.theguardian.com/politics/2019/jul/25/boris-johnson-vows-to-completely-ditch-brexit-backstop; "Boris Johnson says he awaits EU move on Brexit delay after parliament defeat", Alasdair Sandford, accessed October 28, 2019, https://www.euronews.com/2019/10/23/watch-live-boris-johnson-addresses-uk-parliament-after-brexit-vote-defeat; "EU agrees to January 31 Brexit extension", Jacopo Barigazzi and James Randerson, accessed October 28, 2019, https://www.politico.eu/article/eu-agrees-to-january-31-brexit-extension/.

2014, EVI by Amazon and TouchType by Microsoft. Examples of other relevant British AI start-ups and companies are Prowler.io, Graphcore and Improbable.

The UK has been in forefront of producing AI-related strategic documents, reports and studies such as the the AI Sector Deal[214] containing measures to foster AI development in the Country. A few more relevant examples are the House of Commons Robotics and Artificial Intelligence report in 2016[215], Growing the artificial intelligence industry in the UK by Professor Dame Wendy Hall and Jérôme Pesenti in 2017[216], or the Royal Society Machine learning: the power and promise of computers that learn by example also in 2017[217].

The path that the UK decides to take in the future may either make a full member of the EU, an important ally or a powerful competitor. One thing is for certain, if the UK stays it will be a key actor in influencing the EU's AI policy.

§ 4. Italy [218]

Currently, Italy does not possess an AI Strategy, even though there is already a Working Group preparing Italy's AI strategy[219]. In addition, some relatively deep preliminary work has already been carried out, like the Task Force on Artificial Intelligence of the Agency for Digital Italy's "White Paper on Artificial Intelligence at the service of citizens"[220]. The same agency is also coordinating efforts to map the AI ecosystem in Italy with the objective of building synergies and fostering collaboration between AI players in Italy. The initial mapping identified almost 200 entities working on AI related fields in the

[214] *See*, "AI Sector Deal", British Government, accessed June 14, 2019, https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal.
[215] *See*, "Robotics and artificial intelligence", House of Commons Science and Technology Committee, accessed June 14, 2019, https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/145.pdf.
[216] *See*, "Growing the artificial intelligence industry in the UK", accessed June 14, 2019, https://www.gov.uk/government/publications/growing-the-artificial-intelligence-industry-in-the-uk.
[217] *See*, "Machine learning: the power and promise of computers that learn by example", The Royal Society, accessed June 14, 2019, https://royalsociety.org/~/media/policy/projects/machine-learning/publications/machine-learning-report.pdf.
[218] *See,* GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.
[219] *See,* "Intelligenza Artificiale: al via il gruppo di lavoro per una strategia nazionale", Consiglio Nazionale delle Ricerche, accessed June 10, 2019, https://www.cnr.it/it/news/8593/intelligenza-artificiale-al-via-il-gruppo-di-lavoro-per-una-strategia-nazionale.
[220] *See*, "White Paper on Artificial Intelligence at the service of citizens", Task Force on Artificial Intelligence of the Agency for Digital Italy, accessed June 11, 2019, https://ia.italia.it/assets/whitepaper.pdf.

Country[221]. The Observatory on Artificial Intelligence was also established by the AI task force in cooperation with HER - Human Ecosystem Relations to assess the public sentiment about the technology in the Country through the analysis of public conversations about AI in social networks[222].

In the not so distant past, the Member-State was governed by a coalition composed of the far-left Movimento 5 Stelle[223] and the far-right Lega, two, arguably, populist parties. Currently, and after the Lega decided to leave the coalition and seek a snap election, an alternative majority emerged in the Parliament to support a coalition government between the Movimento 5 Stelle and Democratic Party, moving the country closer to the traditional European paradigm. However, this coalition may not be stable, and one cannot stop himself from asking how this instability will affect the country's AI strategy and future development.

§ 5. Spain [224]

The Spanish National AI Strategy was expected to be unveiled in the first semester of 2019, in accordance with the timetable set up by the March 2019, "Estrategia Española de I+D+I en Inteligencia Artificial"[225]. This deadline was not met and, at the time of this writing, there was still no final Spanish AI strategy. Still, and based on the existing documents it is possible to perceive some of the priorities identified by the Spanish State for AI research and development namely: ensuring an organizational environments that foster AI investigation, invest in key strategic sectors, encouraging formation in AI and transfer of knowledge, develop an ecosystem that makes data available for AI development (taking into account the rights of data subjects) and ensuring ethics in AI research and development. Key sectors identified were Public Administration, Education, Security,

---

221 *See*, "Artificial Intelligence Ecosystem", Task Force on Artificial Intelligence of the Agency for Digital Italy, accessed June 11, 2019, https://ia.italia.it/en/ai-in-italy/

222 *See*, "Italian Observatory on Artificial Intelligence", Task Force on Artificial Intelligence of the Agency for Digital Italy, accessed June 11, 2019, https://ia.italia.it/en/ai-observatory/

223 On this point it is interesting to note that the Italian Data Protection Authority recently fined in the amount of €50.000 the party's participatory democracy platform, where they directly consult members for GDPR infringements.

224 *See,* GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

225 *See*, "Estrategia Española de I+D+I en Inteligencia Artificial", Secretaría General de Coordinación de Política Científica del Ministerio de Ciencia, Innovación y Universidades and Grupo de Trabajo en Inteligencia Artificial GTIA, accessed June 12, 2019, http://www.ciencia.gob.es/stfls/MICINN/Ciencia/Ficheros/Estrategia_Inteligencia_Artificial_IDI.pdf.

Tourism and Cultural and Creative Industries, Connected Industries 4.0., Natural Resources and the Natural Environment, Security and Health.

Spanish cities are also playing a central role in fostering digital development in the Country with their smart city initiatives. Madrid and Barcelona achieved a respectable 24th and 28th position in the IESE cities in Motion Index 2019[226]. Barcelona was even considered one of the three smartest cities in the world in a report produced by Philips Lighting and SmartCitiesWorld[227]. Moreover, other, smaller cities in Spain, are running highly complex and interesting smart city initiatives like Santander, Valencia or Málaga.

Spain is also to home to a significant number of technological companies with great future potential. Cabify and Glovo which directly compete with the American company Uber (a leader in the development of autonomous vehicles) are one example. Fon, a company that makes available over 20.000 WiFi hotspots in the world was also founded in Spain. Another example is Red Points that uses AI to provide services to prevent IP infringement.

§ 6. Portugal [228]

The Portuguese economy, albeit relatively small in the European context, is quickly shifting towards emerging technologies, including AI. Portugal will keep hosting Europe's largest technology conference, Web Summit, for the next ten years. Nevertheless, Web Summit is far from the only initiative happening in Portugal, with others as The Lisbon Investment Summit, Dig Publishing and the Singularity University having made their presence known in Portugal in recent years[229]. Portugal has the second largest growth of technology related employment in Europe (6,4%, only behind France's 7.3%)[230] . Companies like Google, BMW, Canon, Cisco, Dell, EPSON, INTEL, PHILIPS and

[226] In the same ranking London achieved the 1st position, Amsterdam 3rd, Paris 4th. European Cities, in general, achieved preeminent positions. *See*, "IESE Cities in Motion Index", IESE Business School, accessed July 10, 2019, https://media.iese.edu/research/pdfs/ST-0509-E.pdf.

[227] "Understanding the challenges and opportunities of smart cities" Philips Lighting and SmartCitiesWorld , accessed June 11, 2019 http://www.lighting.philips.com/main/inspiration/smart-cities/smart-city-trends/smart-cities-world?origin=10_global_en_smartcities_pressrelease___scwnreport_70124000000WUcs

[228] *See*, GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

[229] *See*, "An insider's view of Lisbon's rapidly growing tech scene", Clara Armand-Delille, accessed June 12, 2019, https://tech.eu/free/21951/an-insiders-view-of-lisbons-rapidly-growing-tech-scene/.

[230] *See*, "The State of European Tech 2018", Atomico, accessed June 13, 2019, https://2018.stateofeuropeantech.com/.

Samsumg have offices in Oeiras, Euronext in Porto and in the near future Amazon, Accenture and Bosch in Braga. Companies founded in Portugal like Feedzai, Talkdesk, Unbabel and DefinedCrowd are building a position for themselves on the AI market, others as WeDo, Outsystems, Critical Systems e Primavera Software are incorporating AI in their products.

The Portuguese government plans to use the Golden Visa policy to draw foreign investors in the technology area. Moreover, a plethora of incubators and accelerators as Startup Lisboa, Labs Lisboa, Inovisa, Tec Labs, EIT InnoEnergy and BGI promote by the current technology friend public policies, ensure that there is a healthy environment for start-up development.

§ 6.1. The Portuguese AI Strategy[231]

Portugal made its AI strategy available on June 2019[232]. The writing of the Portuguese strategy has coordinated by University of Porto's Professor Alípio Jorge and is contained within INCoDe.2030, the Portuguese initiative to foster digital skills.

It is not easy to properly evaluate the Portuguese AI strategy. On one hand it is a well-organized document, with some pertinent ideas, that also contains an interesting description of the state of AI development in Portugal, in Europe and even in third countries such as the United States. On the other hand, having a strategy implies that you have a plan to achieve a certain objective/goal. The Portuguese AI strategy has goals but no logical and understandable plan to achieve them.

As an example, section 8 (Strategy) starts by stating that: *i) "The attractiveness of Portugal for knowledge intensive young companies and international production units is high and has conditions to improve; ii) "the development of this ecosystem will motivate the increase of the currently developing innovation levels for a vast number of companies and organizations"; iii) "the research potential in AI and other areas will grow; and iv) "Academia alone and in collaboration with Industry will increase its capacity and develop qualification programs of different levels in AI"*. The scenarios are not unrealistic and the explanation for why each one of them will happen is a likely assessment

231 The Author would like to thank the Ministry of Science, Technology and Higher Education for kindly fulfilling the request for early access to the draft Portuguese AI Strategy.

232 *See,* "AI Portugal 2030: An innovation and growth strategy to foster Artificial Intelligence in Portugal in the European context", INCoDe.2030, accessed July 10, 2019, https://www.incode2030.gov.pt/sites/default/files/incode_aiportugal2030_june19.pdf.

in the context of the development of AI. However, imagining the possible scenarios is just one part of the strategy. A second, and equally important part, entails making sure that the most favourable scenario comes to pass. Indeed, the Portuguese Strategy has some measures that are supposed to accelerate the abovementioned scenarios such as defining regulatory frameworks (including regulatory sandboxes), innovation programmes and support for research and funding on AI related matters. In abstract, none of the ideas contained in this section is bad. Still, their effectiveness may vary greatly depending on how they are implemented, ranging from highly effective to counterproductive. This is where the strategy part fails again, it would be desirable to have a step-by-step plan about how to implement the measures that will help foster AI development in Portugal, instead of some broad strokes of tools that may or may not help. If we want to fund innovative AI companies, will we do it through national funds, European funds or both? What value will be assigned? How will companies be selected? Are there any conditions attached? etc. This type of information is extremely important for a complete strategy and, unfortunately, the Portuguese Strategy is not nearly precise enough on this matter.

Portugal wants to use AI to promote Economic Growth, Scientific Excellence and Human Development and it wants to excel in certain niche areas of AI development as described in the document's chapter about specific actions[233]. But again the specific measures contained in the proposal to get a leadership position are not enough to understand the though process of the authors and/or to evaluate the probability that it works. This is understandable in matters that implicate European or International cooperation (and even than we can, and should, have a strategy to influence European and International decision making on AI), but hardly when the country is only depending on itself to implement a measure. It is not encouraging when the most developed measures are: *i)* the development of centralized administrative data repository managed by Instituto Nacional de Estatística (INE – the Portuguese statistics office) and; *ii)* promoting AI in schools through clubs[234].

[233] Those areas are Natural Language Processing, Real Time decision making with AI, AI for Software Development and AI for Edge-computing. In cooperation relevant players from other countries (companies, academia, research institutes etc.) Portugal also wants to invest in the areas of AI, environment and biodiversity, AI, mobility and autonomous driving, AI and cybersecurity, AI and Health and AI and Industry.
[234] In fact, stating that these are specific measures may be a bit of a stretch. The first one is more developed than most because we actually know who will manage it. The second because it is already happening in plenty of schools. The role of the Atlantic International Research Centre seems more well thought than most of the measures of the Strategy, but that may be due to the fact that this is already a quite well developed project and, thus, since it is highly relevant to this matter and quite targeted, it is easy to give it a role. However, the

§ 7. Finland [235]

Finland is investing strongly on ensuring that its citizens have the necessary AI skills to make the Country a relevant player in the developing landscape. Through initiatives like the free online course, "Elements of AI" and the "AI Challenge", Finland is betting on its workforce as the Country's ace on the AI market. The steering group created by the Minister of Economic Affairs with the aim of designing a proposal for an artificial intelligence program for Finland already submitted two reports regarding artificial intelligence in the Member States. The first one focuses mainly on the economic and ethical side of AI[236] and the second focuses on AI's affects in the labour market. A third, and final, report is expected is expected shortly[237].

Finish AI start-ups have managed to achieve considerable success in the market, with examples like Claned, The Curious AI Company, Kirontech or Wordrive. Furthermore, Finland plans to team up with Sweden and Estonia in AI development. Lastly, Finland will hold the rotating Presidency of the Council of the European Union from 1 July 2019 to 31 December 2019 and thus shall be in a position where it can deeply influence the new EC's strategy on AI and the Digital Single Market[238].

strategy does not develop as much the role of other incredibly important institutions like the International Iberian Nanotechnology Laboratory.

[235] "*See,* GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

[236] *See,* "Finland's Age of Artificial Intelligence: Turning Finland into a leading country in the application of artificial intelligence – Objective and recommendations for measures", Finish Ministry of Economic Affairs and Employment, accessed 10 June 2019, http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf?sequence=1&isAllowed=y.

[237] *See*, "Work in the age of artificial intelligence: Four perspectives on the economy, employment, skills and ethics", Finish Ministry of Economic Affairs and Employment, accessed 10 June 2019, http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160980/TEMjul_21_2018_Work_in_the_age.pdf.

[238] *See*, "Finland's grand AI experiment: Inside Finland's plan to train its population in artificial intelligence", Janosch Delcker, accessed 10 June 2019, https://www.politico.eu/article/finland-one-percent-ai-artificial-intelligence-courses-learning-training/.

§ 8. Estonia [239]

Estonia has been named "the world's most digitally advanced society". In the Country 98% of the interactions between citizens and government are conducted through digital means. Estonia's smart ID cards are used for identification and connection not only with public services such as tax services or healthcare services but also with services provided by the private sector. Due to its geopolitical context Estonia takes cybersecurity particularly seriously and Estonia's public services are designed in a manner that would allow them to keep running even if the country was occupied by a foreign power (they would run from the cloud and possibly from another country). The Country was also a pioneer in the concept of e-Residency which gives foreign entrepreneurs the opportunity to invest in Estonia. In line with what is offered to traditional residents, e-Residents are issued with a digital ID and have full access to Estonia's public e-services. The objective is to enable them to run an EU based business anywhere in the world. E-democracy is a reality in Estonia, where a third of citizens exercise their right to vote online.

Estonia's strategy was initially put in place due to lack of resources to mount a public administration in a standard manner. Due to its success the country kept developing it and now it offers top class public services at a fraction of the cost incurred by other nations. This strategy bore fruit even in the private sector. Companies like Skype and Taxify were founded in Estonia and it boasts a significant number of tech "Unicorns" (companies evaluated at over one thousand million dollars) in comparison with similarly sized Countries[240].

Estonia made its national strategy available in May 2019[241]. As expected, there is a high degree of focus in the public sector as a key actor in fostering innovation. Measures include the creation of Chief Data Officers at least on Ministerial levels, a data science and

---

[239] *See,* GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

[240] *See*, "Successful Stories", e-Estonia, accessed June 12, 2019, https://e-estonia.com/; https://www.cnbc.com/2019/02/08/how-estonia-became-a-digital-society.html; "How a tiny country bordering Russia became one of the most tech-savvy societies in the world", Elizabeth Schulze, accessed June 12, 2019, https://www.cnbc.com/2019/02/08/how-estonia-became-a-digital-society.html; "Welcome to E-stonia, the world's most digitally advanced society", Matt Reynolds, accessed June 12, 2019, https://www.wired.co.uk/article/digital-estonia.

[241] *See,* "Estonia accelerates artificial intelligence development", Estonian Government, accessed July 10, 2019, https://e-estonia.com/estonia-accelerates-artificial-intelligence/; "Estonia's AI Strategy", Estonian Government, accessed July 10, 2019 https://www.riigikantselei.ee/sites/default/files/riigikantselei/strateegiaburoo/eesti_tehisintellekti_kasutuselevotu_eksperdiruhma_aruanne.pdf [In Estonian]

AI collaboration network within the public sector and the development of regulatory sandboxes. Funding of AI development and using AI to safeguard and develop the understand of Estonian language and cultures are also key aspects.

Further, it is likely that the Country will have the first European law specifically regulating liability for AI[242] making it, again, a pioneer in the industry and a case-study for other Member States. From the sociological point of view, Estonia is particularly well-suited for the development of new legal approaches in a digital society because there is a general attitude of readiness to try out new solutions and apply them in practice both in the public and private sectors. All Estonian exceptional advances in e-government, i-voting, electronic identification, e-services etc. became possible thanks to appropriate and timely innovation in legal environment, which in turn became possible thanks to the readiness to adopt new technological phenomena. Estonia may serve as a successful example in which the rest of the EU may draw inspiration. Being at the negotiating table while having a degree of accumulated knowledge and practical experience may also give some leverage to Estonia in the negotiation of AI-related measures.

§ 9. The Netherlands

There is, yet, no AI strategy in the Netherlands. However, a first draft was prepared by the public-private partnership AINED on 6 November 2018[243]. The first draft identifies some concerning trends that the country will try to revert, namely loss of momentum regarding publications an AI research in the Netherlands when compared to other countries, difficulties by SMEs to find the necessary AI expertise to explore the potential of AI and the fact that the Dutch companies developing AI solutions are frequently bought by companies from third-countries creating issues of data ownership. The Dutch Digitalization Strategy[244] also contains strategies and proposed measures to accelerate and strengthen the implementation of digital and emerging technologies in the Netherlands, including AI.

---

[242] *See*, "AI and the Kratt momentum", e-Estonia, accessed June 12, 2019, https://e-estonia.com/ai-and-the-kratt-momentum/.

[243] *See,* GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

[244] *See*, "Dutch Digitalisation Strategy: Getting the Netherlands ready for the digital future", Ministry of Economic Affairs and Climate Policy, accessed June 11, 2019, https://www.government.nl/binaries/government/documents/reports/2018/06/01/dutch-digitalisation-strategy/Dutch+Digitalisation+strategy+def.pdf.

The Netherlands possesses a reasonable share of key companies for the development of AI. Amsterdam, a known "smart city" is home to the European headquarters of such companies as Akzo Nobel, Heineken, ING Group, Cisco Systems, TomTom, Delta Lloyd, Booking.com, Adyen and Philips. Research in Netherlands is also high quality and focuses on ICT challenges, including AI and the city expects to attract a few more due to Brexit[245].

§ 10. Sweden [246]

In 2017, the Swedish government commissioned a report to Vinnova about the impacts of AI in the Country. The report examines Sweden's AI capabilities, the challenges for the Country and key development areas. Moreover, it also contains policy recommendations to boost the Swedish AI economy and examples of the implementation of AI in Swedish projects such as the AI for Breast Cancer Screening coordinated by Karolinska University Hospital or the Aifloo SmartBand developed by Aifloo in cooperation with Skellefteå Municipality that helps senior citizens keep their independence while ensuring that they have access to the help they need in case unforeseen circumstances arise. In 2018, Sweden presented its National AI Strategy that focuses on four pillars: *a)* Education and Training; *b)* Research; *c)* Innovation and use and; *d)* Framework and infrastructure.

The Country plans to use its extensive network of Universities to foster AI research. Programs as the Wallenberg AI, Autonomous Systems and Software Program whose funding amounts to SEK 4 billion (more €350 million) for the its duration (11 years)[247] will play a key role in the Swedish plans for AI development. Additionally, it is planning to use its environmental credentials to expand the national data pool, by luring foreign companies to data centres located in Sweden due to their very low carbon footprint (and cheap price) in comparison with some competitors like the United States or China[248].

---

[245] "Artificial Intelligence in Amsterdam, the City of Freedom", Simona Nickman, accessed June 11, 2019, https://medium.com/cityai/artificial-intelligence-in-amsterdam-the-city-of-freedom-83406e866e7e

[246] *See,* GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

[247] *See*, "Wallenberg AI, Autonomous Systems and Software Program (WASP)", Linköping University, accessed June 10, 2019, https://liu.se/en/research/wallenberg-artificial-intelligence-autonomous-systems-and-software-program.

[248] *See*, "Sweden set to become global leader in artificial intelligence", The Swedish Trade and Investment Council, accessed June 10, 2019, https://www.business-sweden.se/en/Invest/industries/Data-Centers-By-

§ 11. Belgium [249]

The scenario in Belgium is different from the one in other central European Countries. It has yet to catch up to Germany, or even to the Netherlands, on the size of its industry, specifically where it is related to AI. Belgium's multicultural characteristics make it so that algorithmic bias could potentially hit this country particularly hard.

Even though the legal, pollical and strategic work in Belgium is not very advanced in regard to AI implementation, we must keep in mind the Country's heavy influence on European law-making. With both the European Council, the Council of the European Union and the European Commission headquartered in Brussels, and most of the European Parliament's work also carried out in the city (albeit its headquarters are in Strasbourg), Belgian solutions and opinions are in a prime position to influence the European process. Furthermore, the private AI sector in Belgium appears to be going strong, even with humble public initiatives behind it. A study by Microsoft reported that over the past decade there were 14 transactions involving AI companies with a value of €110 Million in the Countries of Brussels and Luxembourg, a larger value than in the Netherlands or Italy[250].

At the time of this writing, no official Belgian AI strategy existed. However, a group of 40 experts (the AI4Belgium coalition) recently proposed a draft strategy with the backing of Alexander De Croo and Philippe De Backer (government officials).This report, focusing on education, skill building, building, ensuring that data is shared, fostering innovation and supporting adoption of AI-enabled technologies in both public and private sector is posed to be a key step for the official Belgian AI strategy[251].

Sweden/news-and-downloads/investment-news/sweden-set-to-become-global-leader-in-artificial-intelligence/.

[249] *See,* GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.

[250] The countries were analysed jointly in the report. *See*, "Artificial Intelligence in Belgium and Luxembourg: How 277 major European companies benefit from AI", Microsoft, accessed June 10, 2019, https://info.microsoft.com/WE-DIGTRNS-CNTNT-FY19-09Sep-27-ArtificialIntelligenceinBelgium-MGC0003166_01Registration-ForminBody.html?wt.mc_id=AID732606_QSG_280351.

[251] "AI4Belgium", AI4Belgium coalition, July 11, 2019, https://www.ai4belgium.be/wp-content/uploads/2019/04/report_en.pdf

PART II – CURRENT EU LEGAL FRAMEWORK AND CHALLENGES

Chapter I – Greedy computers: machine learning and data processing

§ 1. The GDPR

Any effort to discuss AI regulation in the European Union (and, in fact, in the world) cannot ignore that there is already a comprehensive and harmonised regulation for one of the most important aspects of AI development: data protection. Regulation 2016/679/EU 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, generally known as the General Data Protection Regulation or GDPR. With the GDPR, the European Union positioned itself as the world's data protection police and standard-setter. The GDPR contains a comprehensive regulation of data protection matters that will apply to most data processing operations[252]. Non-compliance with the GDPR carries the risk of hefty fines along with civil liability and potentially other sanctions. Furthermore, the European Union is now including data protection matters in trade deals established with other economic blocks. One notable example is the trade deal between the EU and Japan that also included data protection in the negotiations and ended up with mutual recognition of equivalent levels of protection.

The GDPR is undoubtedly not the regulatory bogeyman that some adepts of a more *laissez-faire* approach to data protection want to make it. That is to say, the GDPR does not outright forbid most data processing operations, but it does set forth relatively demanding requirements for those data protection operations. Said requirements are stricter on data protection operations using special categories of personal data or that may, for example, result in a high risk to the rights and freedoms of natural persons. It is built upon an idea that a person owns his/her data and thus, it also contains, a plethora of rights that may be exercised by the data subject. In addition, as an uninformed data subject is not

[252] Where no special regime applies such as the 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 (hereinafter, "e-Privacy Directive") and Directive 2016/680/EU of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA. Even then, some GDPR rules may be applicable by remission, such is the case of e-Privacy Directive which refers to the definitions contained in the GDPR in its Article 2.

in any position to make decisions regarding the data that is owned by said data subject, the GDPR also contains mandatory information that must be provided to the data subject.

The sanctions for not complying, as abovementioned, are steep and may arise from various sources, including administrative fines and civil liability. It is expected that under the GDPR, national data protection supervisory authorities take an active role in enforcement and are granted the means (including technical means and human resources) to do. While they had somewhat of a cold start in doing so[253], data protection supervisory authorities across Europe are starting to levy considerable fines to non-compliant data controllers[254]. In addition, other penalties may be enshrined into law by Member States, including potential criminal liability and restrictions to data processing operations.

ML is heavily dependent on data, including but not limited to personal data, and therefore ensuring that said data is available for development is key for success in this area. Of course, if we confront the rules contained within the GDPR with what we already discussed about machine learning development, it is possible to identify some possible conflicts. However, it is important to understand if they are insuperable issues that may completely block or, at least, severely hinder the development of ML in the EU (the so-called "chilling effect") or if they are, in contrast, conflicts that can be solved through good planning, practices and corporate governance. In the chapter below, we will analyse the most relevant of those conflicts and suggest potential solutions. We will focus on general

[253] Issues such as the lack of national legislation giving them the adequate resources to answer data subjects' complaints or investigate potential data protection violations may have contributed to this slow start. However, those problems seem to have been mostly dealt with.

[254] Four examples: ICO fined British Airways £183.39 Million and Marriott £99 Million, while CNIL fined Google €50 Million, the Bulgarian Data Protection agency levied a fine of over €2.5 Million against Bulgaria's National Revenue agency after the leak of more than 4 million Bulgarian citizens' data. *See,* "Intention to fine British Airways £183.39m under GDPR for data breach", ICO, accessed September 6, 2019, https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/07/ico-announces-intention-to-fine-british-airways/; "Statement: Intention to fine Marriott International, Inc more than £99 million under GDPR for data breach", ICO, accessed September 6, 2019, https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/07/statement-intention-to-fine-marriott-international-inc-more-than-99-million-under-gdpr-for-data-breach/; "The CNIL's restricted committee imposes a financial penalty of 50 Million euros against GOOGLE LLC, CNIL, accessed September 6, 2019, https://www.cnil.fr/en/cnils-restricted-committee-imposes-financial-penalty-50-million-euros-against-google-llc; "Bulgaria's tax agency fined $3 million over data breach, will appeal", Reuters, accessed September 6, 2019, https://www.reuters.com/article/us-bulgaria-cybersecurity-fine/bulgarias-tax-agency-fined-3-million-over-data-breach-will-appeal-idUSKCN1VJ0YY.

issues, as specific issues can always appear in specific situations, but they can only be adequately assessed on a case-by-case basis[255]/[256].

[255] For example, ML may be quite troublesome when performing a Data Protection Impact Assessment. However, the analysis of the eventual challenges is quite tied to the specific application of ML that we want to make in case. It would be possible to think about and create a few standard scenarios and it would certainly be interesting to do so, but both space and time are limited, and this may not be the most adequate place for that exercise.

[256] See, The first steps of a revolution with a set date (25 May 2018): the "new" General Data Protection regime", Pedro Madeira Froufe, accessed June 17, 2019, https://officialblogofunio.com/2018/05/25/the-first-steps-of-a-revolution-with-a-set-date-25-may-2018-the-new-general-data-protection-regime/; "Implications of the declaration of invalidity of the Directive 2006/24 on the retention of personal data (metadata) in the EU Member States: an approach to the judgment Tele 2 of 21 December 2016", Alessandra Silveira and Pedro Miguel Freitas, accessed June 8, 2019, https://officialblogofunio.com/2017/01/22/implications-of-the-declaration-of-invalidity-of-the-directive-200624-on-the-retention-of-personal-data-metadata-in-the-member-states-of-the-eu-an-approach-to-the-judgment-tele-2-of-21-december-20/; "Data Protection Officer according to GDPR", André Mendes Costa, accessed July 9, 2019, https://officialblogofunio.com/2017/06/13/2014/; "Data Protection, Data Transfers, and International Agreements: the CJEU's Opinion 1/15", Christopher Kuner, accessed July, 2019, https://verfassungsblog.de/data-protection-data-transfers-and-international-agreements-the-cjeus-opinion-115/; "The EU General Data Protection Regulation: Powerful Tool for Data Subjects?", Enrico Peuker, accessed July 9, 2019, https://verfassungsblog.de/the-eu-general-data-protection-regulation-powerful-tool-for-data-subjects/; Alessandra Silveira e Pedro Miguel Freitas, *"The recent jurisprudence of the CJEU on personal data .on: implications for criminal investigation in Portugal"*, UNIO – *EU Law Journal* 3,2 (2017): 45-56; "The Three Laws of Robotics in the Age of Big Data", Jack M. Balkin, accessed July 10, 201~9, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2890965; "Exploring or Exploiting? Social and Ethical Implications of Autonomous Experimentation in AI", Sarah Bird et al., accessed June 3, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2846909; "Humans Forget, Machines Remember: Artificial Intelligence and the Right to Be Forgotten", Eduard Fosch Villaronga, Peter Kieseberg and Tiffany Li, accessed June 9, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3018186; Efren Díaz Díaz, *"The new European Union General Regulation on Data Protection and the legal consequences for institutions"*, Church, Communication and Culture 1,1 (2016): 206-239; Paul B. Lambert, *Understanding the New European Data Protection Rules* (New York: CRC Press, 2018):197ff.; "GDPR and the Challenges of Digital Memory", Kiran Wattamwar, accessed March 15, 2019, http://sitn.hms.harvard.edu/flash/2018/gdpr-challenges-digital-memory/; "Trends shaping AI in business and main changes in the legal landscape", Ana Landeta and Felipe Debasa, accessed February 25, 2019, https://officialblogofunio.com/2019/02/20/trends-shaping-ai-in-business-and-main-changes-in-the-legal-landscape/; Francisco Pacheco de Andrade, Pedro Miguel Freitas and Teresa Coelho Moreira, "Data Protection and Biometric Data: European Union Legislation", *in. Biometric Security and Privacy: Opportunities & Challenges in The Big Data Era*, Richard Jiant et al. ed., (Switzerland: Springer: 2017), 413-421; Francisco Pacheco de Andrade and Teresa Coelho Moreira, "Personal Data and Surveillance: The Danger of the "Homo Conectus"", *Intelligent Environments 2016: Workshop Proceedings of the 12th International Conference on Intelligent Environments*, in Paulo Novais and Shin'ichi Konomi (Amsterdam: IOS Press, 2016), 115-124; Graça Canto Moniz, *"Finally: a coherent framework for the extraterritorial scope of EU data protection law - the end of the linguistic conundrum of Article 3(2) of the GDPR"* UNIO EU Law Journal.4,2 (2018): 105-116; Pedro Miguel Freitas, *"The General Data Protection Regulation: an overview of the penalties' provisions from a Portuguese standpoint"* EU Law Journal.4,2 (2018): 99-104; Chris Holder et al., *"Robotics and law: Key legal and regulatory implications of the robotics age (Part I of II)"* Computer Law & Security Review 32,3 (2016): 383-402; "Google v. CNIL: Is a new landmark judgment for personal data protection on the horizon?", Alessandra Silveira and Tiago Sérgio Cabral, accessed September 7, 2019, https://officialblogofunio.com/2019/09/10/editorial-of-september-2019/.

§ 1.1. A Quick Look into the Relationship between AI and the General Principles of Data Protection

While in the sections below we will identify and consider possible solutions for the main issues that may arise when developing and deploying AI under the GDPR, a quick look into the relationship between AI and the general principles of data protection should, certainly, be in store.

In accordance to the principles of lawfulness, fairness and transparency personal data shall be processed in a lawful, fair and transparent manner in relation to the data subject. The key aspect here, at least in what concerns the current development of AI, is really transparency. Issues such as black box algorithms, non-understandable AI and the gap in knowledge between the common citizen and specialists can make compliance with these principles quite difficult. If a citizen cannot understand if, how and why data is being processed if is not possible to assess if the processing operations are being carried in accordance to the law or in a fair manner. Of course, transparency is also not being realised in any sense of the concept. Below, in the sections about explainability and the right to information the issue will be dealt with in a more profound manner, and the solutions there can help with general problems arising with these principles.

Purpose limitation can also only be realised through transparent algorithms, as you cannot limit something you do not understand. Therefore, we refer to the same sections as above.

Data accuracy is other challenge, but one that developers should also be keen to tackle, as incorrect data will produce results. Ensuring that proper and accurate datasets are available for developers is key in this area. The HLG and the EC have been stressing this aspect in the public strategic documentation related to AI development, so the issue is identified, now it is matter of constructing or making available through other means said datasets. Having, adequate remedies available where the principle of data accuracy is not complied with can also help, at least by fostering citizen's trust. We develop the issue on our sections on erasure and rectification[257].

[257] The Ibero-American Data Protection Network, which includes the Spanish AEPD, Autoridad Catalana de Protección de Datos and Agencia Vasca de Protección de Datos, and the Portuguese CNPD issued Guidelines on the processing of personal data in AI. Although they are not adapted to the GDPR, since most members are not European, they do provide some interesting pointers on matters such as the actors involved on the processing of personal data for AI. Issues as data quality, need to conduct a (D)PIA when the legal requirements are met, privacy, security and ethics by design, accountability (for damages caused by

Both data minimization and storage limitation will be highly dependent on the specific application of AI. However, the core issue, again, circles back to transparency and erasure (in the broad sense and not specifically the right to erasure). Transparency to know what data is being collected and erasure to suppress the data from the system and to know when it is really erased or anonymized. The sections on erasure, information and explainability will offer some insight. Accountability will be transversal and appear frequently in our chapter about the GDPR.

Last, the principle of integrity and confidentiality can only be complied through technical and organisational measures. We will address the issue lightly in our conclusions and policy proposals at the end of Thesis. Nevertheless, we would note that the industry contribution will extremely important by developing technological solutions and cooperating. Development and voluntary adherence to industry standards is one significant example of industry cooperation that may even, by itself, allow for compliance with this principle.

§ 1.2. Who is the Data Controller?

The GDPR defines data controller as the *"the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data"*[258]. In the current state of affairs, and taking into account the legal definition contained within the GDPR, there are no doubts that an AI-enabled machine cannot be the data controller for data protection purposes. However, there some aspects that should be addressed regarding AI and the concept of data controller under the GDPR.

If the AI cannot be considered as the data controller, the most likely candidate to be the data controller is the natural or legal person who deploys the AI. The natural or legal person who deploys the AI will, with all probability, define the, at least, the purposes and, and most likely, the essential means of processing. Now, some issues may arise when the AI is deployed by a person who is not the developer or deployed in partnership with the developer. In line with what happened in *Fashion Id.* and *Wirtschaftsakademie* this type

---

the processing of personal data), and transparency are some of the concerns shared by the Ibero-American data protection Regulators. *See*, Ibero-American Data Protection Network, *General Recommendations for the Processing of Personal Data in Artificial Intelligence* (Juárez: Mexico, 2019).

[258] Additionally, *"where the purposes and means of such processing are determined by Union or Member State law, the controller or the specific criteria for its nomination may be provided for by Union or Member State law".*

of situations may give rise to limited joint controllerships, which will have to be dealt in a case-by-case basis as the possibilities are endless[259].

A second issue may arise from the data controller "losing control" over the definition of the essential means of processing. According to the WP29 essential means include "*which data shall be processed*?", "*for how long shall they be processed*?" and "*who shall have access to them*?"[260]. It is not far-fetched to imagine a sufficiently advanced system deciding that, to achieve the purposes given to it, it must collect and process some types of data but not others, or deciding that it only needs certain types of data for a limited period of time and after that they can be deleted. If the data controller is not aware of this due, for example, a black box algorithm, it may be unable to comply with its legal obligations (for example providing information on data storage limitation).

Moreover, an entity which was hired as a processor may find itself in position of an unwilling data controller, due to implementing AI-enabled software and/or hardware, that may make decisions that should be reserved to the (original) data controller[261]. This is an additional reason why opaque and non-understandable AI may harm not only the final data subject/consumer but also the entities who deploy it. It is important for entities to know what is happening so that they do not fall in regulatory pitfalls due to negligence.

Now we did say that it is impossible under current rules for AI to be considered as a data controller. But that is more due to the fact that it is neither a natural nor a legal person, as there is no other disposition in the GDPR that forbids it. If, either through national or European law, AI was given a legal status akin to companies, it could, in theory, be considered as the data controller for certain data protection operations. At the current stage of development, where we only developed (quite) Narrow AI this does not seem to be desirable as we will explain below (*see* our chapter on the Product Liability Directive). For data protection, it could create serious impairments to the enforcement of our current data protection rules. Immediately, the GDPR's administrative fines are not designed to work on AI-enabled machines (and we would argue that other sanctions designed by Member States under Article 84 GDPR are in the same situation). But there is more, and more important, AI-enabled machines at the current stage of development would not be

[259] *See*,Judgment of the ECJ of 29 July 2019, *Fashion Id.*, Case C-40/17, ECLI:EU:C:2019:629; Judgment of the ECJ of 5 June 2019, *Wirtschaftsakademie*, Case C-210/16, ECLI:EU:C:2018:388.
[260] Opinion 1/2010 on the concepts of "controller" and "processor", WP29, accessed June 10, 2019, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2010/wp169_en.pdf
[261] This in accordance with the WP29's interpretation which considers that decisions such as "*which hardware or software shall be used*" may be delegated to controllers.

able to collaborate with supervisory authorities, ensure the exercise of the data subject's rights or, conduct a Data Protection Impact Assessment or, where necessary, nominate a DPO.

§ 1.3. The Right to Erasure (also known as the Right to be Forgotten)

§ 1.3.1. Under Directive Directive 95/46/EC and the Google Spain Judgment

Back in 2014, the ECJ issued its judgment on the landmark case of *Google Spain v. Agencia Española de Protección de Datos* and with it brought new breath to discussions around data protection reform[262]. This dispute arose from a complaint submitted to the AEPD (the Spanish data protection supervisory authority) by a Spanish citizen who requested the removal of the information related to past social security debts from the newspaper "La Vanguardia" and from Google Search results[263]. The issue ended being brought before the ECJ, with the Court deciding that the search engine's activity was a type of data processing and, in this context, Google should be classified as a data controller under Directive 95/46/EC. The Court further considered that provisions of Directive 95/46/EC should be interpreted as meaning that Google was required to *"remove from the list of results displayed following a search made on the basis of a person's name links to web pages, published by third parties and containing information relating to that person, also in a case where that name or information is not erased beforehand or simultaneously from those web pages, and even, as the case may be, when its publication in itself on those pages is lawful".*

In addition, the Court's position is that the data subject's rights override the economic rights of the data controller (the right to make the information available and to guarantee the accuracy of its search engine) and the public's interest in having access to the abovementioned information. This argument should hold unless a pressing public interest in having access to the information is present.

---

262 *See,* Judgment of the ECJ of 13 May 2014, *Google Spain v. Agencia Española de Protección de Datos*, Case C-131/12, ECLI:EU:C:2014:317.

263 The ECJ was only asked about the processing of data by Google and not by the newspaper. In fact, the processing of data by a search engine can be unlawful while the website from where the data originates is processing the data lawfully, as the legal basis will, likely be different. In fact, the webpage at the root of the controversy is still online and available. *See*, "Edición del lunes, 09 marzo 1998", La Vanguardia, accessed in June 4, http://hemeroteca.lavanguardia.com/preview/2013/02/27/pagina-13/33837533/pdf.html.

We will not delve deeply into the Google Spain judgment, but there are three facts to which we would like to call the reader's attention. First, the judgment does not appear in vacuum. Directive 95/46/EC contained dispositions that granted the right to have data that did not comply with the Directive – particularly incomplete or inaccurate data –, rectified, erased or blocked to the data subject. Moreover, the data subject had (and still has) the right to object to data processing based on his/her particular situation. Principles such as accuracy, purpose limitation, adequacy and fair and limited personal data protection also precede the current legal framework and already existed in Directive 95/46/EC. Thus, the real and most pressing innovation under Google Spain was related to the fact that it clarified that search engine's activities entailed personal data processing and, in this context, search engines were data controllers and had to comply with the rights there enshrined.

Second, the nature of the right arising from Google Spain is not exactly clear and, in fact, may even be deemed as controversial[264/265]. We cannot affirm with certainty that it is a subset of the right to erasure because, in fact, there is no demand for the erasure of the data, even from Google's servers. The requirement is just to suppress certain hyperlinks from the public results list if certain queries are used in the search. It is, in our opinion, "targeted delisting". Thereby, it would be more adequate to tell that there is a restriction of the processing for a certain purpose, in this case to display a public list of results, previously indexed through technical means such as web crawlers, and organized by an algorithm in accordance with perceived relevance to the user base, when a specific query is used. In fact, from the judgment itself it does not result that the search engine cannot index the website as before and display the result if the query differs from the name of the data subject.

---

[264] It is true that, in other works, we have called it the right to non-indexation. However, indexation does happen. The user may only ask for the results containing personal data related to him/her to be suppressed after the list of results is public, detectable and detected. Therefore, calling it non-indexation is not entirely accurate. An example of pure non-indexation would be, for example, the orders given to crawlers through the ROBOT meta tags and robots.text files, to not index the page and not follow hyperlinks in said page, or the privacy definitions that can be selected on certain social network profiles to achieve the same result.
One alternative that could be considered would be calling it deindexation, since it would solve the problem of the existence of an initial indexation. Still, even if deindexation is accurate as far as the order of the procedure goes, it is not in relation to the limitations of it. Google Spain does not give the data subject the right to be completely erased from the index, it gives him/her the right for the hyperlinks concerning his/hers personal data to not be displayed in the public list of results, when the search is based on the person's name or, we would argue, other personal data. Therefore, the company could, in theory, crawl the webpage as normal, index, build a list and, just before public display, apply a filter that would suppress the hyperlinks that have been restricted in accordance with the judgment. Furthermore, the results will still appear if the search query is different from the personal data related to the data subject and whose processing was limited. CNPD has the right idea when it considers that this is a right to "delisting", but it is still not one hundred percent accurate. Delisting would imply that the result is completely removed from the list independently of the search query used. The supervisory authority itself does not think that this is the case, the results can still appear if a difference search query is used and, thus, the term "target delisting" should be used.
That is why Google Spain should be arguably, read as containing a right to "target delisting from the public list of results when specific queries are used". *See*, "Protecting our personal data in the 21st century: why the new EU legal framework matters", Rita de Sousa Costa and Tiago Sérgio Cabral, accessed in June 3, 2019, https://officialblogofunio.com/2016/06/20/protecting-our-personal-data-in-the-21st-century-why-the-new-eu-eu-legal-framework-matters/#_edn13; "Best Practices for Setting Up Meta Robots Tags and Robots.txt", Sergey Grybniak, accessed in June 4, 2019, https://www.searchenginejournal.com/best-practices-setting-meta-robots-tags-robots-txt/188655/#close; "Parecer n.º 20/2018 da CNPD, relativo à Proposta de Lei 120/XVIII", accessed in June 2, , 2019, 34v, http://app.parlamento.pt/webutils/docs/doc.pdf?path=6148523063446f764c324679626d56304c334e706447567a4c31684a53556c4d5a5763765130394e4c7a464451554e4554563765247396a6457316c626e527663306c7561574e7059585270646d4644623232317063334e686279396a5a57593359544d794f4330325a44526c4c54526c4e546b74596a41304e4331694e54426d4f5449314d6a64684d7a45756347526d&fich=cef7a328-6d4e-4e59-b044-b50f92527a31.pdf&Inline=true.
[265] *See*, Ana Azurmendi, "Spain: The Right to Be Forgotten", *in* Wolf J. Schünemann / Max-Otto Baumann, *Privacy, Data Protection and Cybersecurity in Europe* (Zug, Springer International Publishing AG, 2017), 17-30. Alessandro Mantelero, *"The EU Proposal for a General Data Protection Regulation and the roots of the 'right to be forgotten'"*, Computer Law & Security Review 29,3 (2013): 229-235; Ioannis Iglezakis, *"The Right to Be Forgotten: A New Digital Right for Cyberspace"* Cyberlaw 1,3 (2017): 67-79.

Third, the right of erasure is significantly more developed in the GDPR when compared to Directive 95/46/EC and the Google Spain judgment. In fact, the European legislator seems to have used Google Spain as a stepping stone to build upon and create a more detailed and comprehensive framework[266].

§ 1.3.2. Under the GDPR

Article 17 of the GDPR grants the data subject the right to obtain, from the controller, the erasure of personal data concerning him/her. In this context, the controller must guarantee that, if the legal requirements are met, the data is erased without undue delay. The development level of Article 17 far surpasses what previously existed in the EU's legal framework. It regulates in which situations the data subject has the right to erasure and exceptions to the general rule. Under this norm the data subject has the right to have his/her information erased when: *i) "the personal data are no longer necessary in relation to the purposes for which they were collected or otherwise processed"*; *ii) the data subject withdraws consent on which the processing is based according to point (a) of Article 6(1), or point (a) of Article 9(2), and where there is no other legal ground for the processing; iii) the data subject objects to the processing pursuant to Article 21(1) and there are no overriding legitimate grounds for the processing, or the data subject objects to the processing pursuant to Article 21(2); iv) the personal data have been unlawfully processed; v) the personal data have to be erased for compliance with a legal obligation in Union or Member State law to which the controller is subject or; the personal data have been collected in relation to the offer of information society services referred to in Article 8(1)"* (offered to children).

Obviously, even in these situations there are relatively broad limitations to the right to erasure, namely if the processing is necessary to: *i)* to guarantee freedom of expression and information; *ii)* comply with a legal obligation to which the data controller is bound; *iii)* for reasons of public interest related to the area of health[267]; *v)* for archiving purposes

[266] The innovative intervention of the ECJ was not lost under the GDPR. Just recently, while we were finishing this Thesis the Court issued two landmark judgments: *Google v. CNIL* and *Planet49*. In *Google v. CNIL* the ECJ restricted the territorial scope of the right to target delisting to the territory of the EU and in *Planet49* clarified that consent for non-essential cookies under the e-Privacy Directive was still needed even if they did not collect personal data, and pre-checked boxes cannot be considered as valid consent. *See*, "Google v. CNIL: Is a new landmark judgment for personal data protection on the horizon?", Alessandra Silveira and Tiago Sérgio Cabral, accessed October 7, 2019, https://officialblogofunio.com/2019/09/10/editorial-of-september-2019/; Judgment of the ECJ of 24 September 2019, *Google v. CNIL*, Case C-507/17, ECLI:EU:C:2019:772; Judgment of the ECJ of 1 October 2019, *Planet49*, Case C-673/17, ECLI:EU:C:2019:801.

[267] Specifically when it is necessary for "*the purposes of preventive or occupational medicine, for the assessment of the working capacity of the employee, medical diagnosis, the provision of health or social care or treatment or the management of*

in the public interest, scientific research historical research, when the erasure of the data is likely to render impossible or seriously impair complications to the aims of the processing activities or; *vi)* for the establishment, exercise or defence of legal claims.

Moreover, the new GDPR updated data privacy framework contains a set of measures intended to ease the burden over the data subject when exercising his/her rights. This can be especially relevant when data is processed by more than one entity. If the data is processed by joint controllers, they will have to contractually regulate their obligations regarding the exercise of data subject rights and make the essence of this arrangement available to data subject[268]. With this in mind, the data subject maintains the right to address any of the join controllers when exercising his/her rights. Even if contractually it is not responsible for the task, the data controller must still ensure compliance with the data subject's request (if there is a legal basis for it).

When the data processing is carried by a data processor on behalf of the controller, the contract (or other legal act under EU or Member State law) governing the relationship between processor and controller must contain provisions by which the processor is obliged to assist the controller in complying with rights of the data subject's rights.

The controller is burdened with the obligation to communicate the request to each recipient of the personal data, unless such endeavour proves to be impossible or involves a disproportionate effort (Article 19 GDPR). If the original controller made the data public, it will have to implement reasonable measures to (attempt to) inform other controllers who are processing the previously publicly available data of the request to erase the personal data (Article 17(2) GDPR)[269].

---

health or social care systems and services on the basis of Union or Member State law or pursuant to contract with a health professional" or "*for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health or ensuring high standards of quality and safety of health care and of medicinal products or medical devices, on the basis of Union or Member State law which provides for suitable and specific measures to safeguard the rights and freedoms of the data subject, in particular professional secrecy*". In this context, an additional safeguard, is that data can only be processed either by professionals that are under the obligation of professional secrecy or other person who is bound by adequate obligations of secrecy (which can arise from a contract with, for example, the healthcare institution).

[268] For what is worth, what constitutes the essence of the arrangements is not clarified in the GDPR, which can create some practical difficulties for data subjects and data processors.

[269] *See*, Information Commissioner's Office, *Guide to the General Data Protection Regulation*, (London: ICO, May 2019): 119ff.; "Opinion 1/2010 on the concepts of "controller" and "processor", WP29, accessed June 10, 2019, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2010/wp169_en.pdf; "A Practical Guide to Controller-Processor Contracts", DPC, accessed June 10, 2019, https://www.dataprotection.ie/sites/default/files/uploads/2019-06/190624%20Practical%20Guide%20to%20Controller-Processor%20Contracts.pdf; Paul Voigt and Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Cham: Springer: 2018):156-164.

Obviously, even in "normal" situations, the correct application of the right to erasure in the GDPR still leaves quite some room for doubts and discussions. For example, when is the exact moment when personal data stops being needed for the purposes for which it was collected? When do the legitimate interests of the controller override the rights and freedoms of the data subject[270]? What data is needed to comply with legal obligations? Frequently laws only lay down an obligation that the data subject needs to comply with, such as not discriminating when selecting new employees, but it does not specify what data must be kept in order to prove compliance. What are the limits of freedom of expression and information?

The ECJ recently addressed this last issue on its *Buivids* judgment in which it considered that the concept of journalist purposes must be interpreted broadly, and it is not restricted to media undertakings. In fact, according to the Court, it should be applied to situations when the sole object of the data processing is "*the disclosure of information, opinions or ideas to the public*[271]".

Back to the doubts and discussions around the right to erasure: until when is data needed to for the establishment, exercise or defence of legal claims?[272] Probably until it is

---

[270] We will not make a deep analysis of the balancing test between the legitimate interests of the controller and the rights and freedoms of the data subject. However, we should take into account that it should be conducted in an impartial manner and, in Google Spain, the European Court of Justice seems to consider that, when in doubt, the balancing test will favour the data subject and not the controller.

[271] See, Judgment of the ECJ of 14 February 2019, Sergejs Buivids v. Datu valsts inspekcija, Case C–345/17, ECLI:EU:C:2019:122. Some Authors argue that, this judgment, even if the broadens the derogation regarding freedom of expression and information, actually restricts the exception of purely personal or household activity. With all due respect, we cannot agree with such a conclusion. The case at hand clearly exceeded what is personal or household activity. In fact, the central question regarded the recording of an interaction where all parties are within an official setting and clearly not an household setting. The police officers were going about their duties in the police station, while Mr. Buivids was delivering a statement in the context of administrative proceedings which had been brought against him. The intention of Mr. Buivids of bringing "*attention of society something which he considered to constitute unlawful conduct on the part of the police* clearly goes further than simple household or personal activity. In this context, it is difficult to say that just because the personal/household activity exception is not applicable here, it will not be applicable to the creator who is recording a video in a public park or to the mother that takes pictures at her children's recital and then publishes them online. *See*, "European Data Protection and Freedom of Expression after Buivids: an Increasingly Significant Tension", David Erdos, accessed August 20, 2019, https://europeanlawblog.eu/2019/02/21/european-data-protection-and-freedom-of-expression-after-buivids-an-increasingly-significant-tension/; Alessandra Silveira e Pedro Madeira Froufe, *"From the Internal Market to the citizenship of rights: the protection of personal data as the jus-fundamental identity question of our times",* UNIO EU Law Journal.4,2 (2018): 3-17; Elena Esposito, *"Algorithmic memory and the right to be forgotten on the web"* Big Data & Society (2017): 1-11.

[272] The broader question is until when does data have value? Data will generally have immediate value and potential value that may or not be realised. Our examples regarding legal claims are examples of data that has potential value, but that value may never be realised if a legal claim is never brought against the data controller. The same could be said about the data in analysis in Google Spain. The Court considered that the data did not have any more immediate value for Google, but if Mr Mario Costeja González became a public figure, it is arguable that it could gain enough value to reverse the logic behind the judgment. Therefore, the data still had potential value. The question is, should data that only has potential value be stored or deleted?

not possible to theoretically pursue said claim. After all, if I entered into a contract 19 years ago and the other party decides to claim that I did not comply with the contract and if prescription did not occur in accordance with national law, I will need data related to the contract to prove compliance[273]. If I ran a store with CCTV and someone, one year from now, claimed to have slipped on my wet and unclearly marked floor, having the CCTV footage would be useful to prove that the person did not slip, and my floor was not even wet.

Still, the line has probably to be drawn somewhere. On the first situation, not all data is necessarily needed. Let us imagine that I was hired to take photographs at a wedding. It would be excessive to keep all photographs to prove that the contract was completed successfully. Asking the other party to sign a declaration demonstrating compliance might work in some instances. Then it is possible to keep said declaration instead of other, more sensitive personal data. At the end of the day the burden of proof lies on the claimant, so in certain cases if someone is only pursuing a claim decades after the occurrence, meeting the burden of proof would be very difficult and the act of only claiming non-compliance a couple of decades after might also hold some significance. Of course, this is a case-by-case analysis and in some cases keeping everything is absolutely needed (for selling objects that will be used for a long time or for the provisions of services that repeat frequently for a significant period or that keep affecting a person for that period, for example). For the CCTV example, the personal data would be useful but is probably not needed, since, again, the burden of proof rests with the claimant.

While in theory it may seem like an easy exercise, this "drawing the line" is not as easy as it sounds. More data helps you build a clearer picture and that has its advantages and disadvantages. More transparency, more information, but less privacy and irrelevant information or outdated/false information mixed with pertinent one. As we will see, this is a challenge for ML, which generally "likes" to have data and, once the original input happens, has some trouble "forgetting" it.

What level of value does data need to have to justify keeping it? *See*, Meg Leta Ambrose, *"It's About Time: Privacy, Information Life Cycles, and the Right to be Forgotten"* Stanford Technology Law Review 16,2 (2013): 101-154.

[273] This solution is contained in the Portuguese data protection law, approved to enable the application of the GDPR.

§ 1.3.3. The Right to Erasure and AI

§ 1.3.3.1. Deleting Data in Computers

The right to be erasure under the GDPR is written in a manner suitable for non-data intensive environments, but not necessarily to very large databases, let alone AI. Data input and data deletion in ML does not work as it would work for the human mind. It is not just about remembering and then forgetting. Moreover, it does not work in a manner where you could simply locate the information and delete it, as you would do on your computer (and even on your computer things are not as simple as that, and we will go there in a moment) [274].

Indeed, even on large, big data adapted databases, the right to erasure may be difficult and/or costly. First, these types of databases have plenty of build-in mechanisms and failsafe measures such as durability measures to avoid data loss and corruption: automatic backups, the ability to rollback to a previous version, control mechanisms etc. (those mechanisms are actually contained in Article 32 of the GDPR). Meaning data is frequently stored in more than one place and it may be difficult to identify and delete all the "copies". But, by itself, it is possible to overcome this issue. Privacy by design is one of the principles of the GDPR and, obviously, mechanisms have to be adapted to its provisions. It may not be easy to find the data, but it is not impossible. Still, if we are talking about cost there is one further challenge that must be addressed: in modern computer science deleting data rarely is what most people would think. From the most advanced databases to our personal computer, by default, deleting a file generally does not really delete the file from the computer. Space where the file was previously saved is just marked as empty and removed from the index/search functions. It is not immediately overridden with zeros. Eventually, it will be overridden by a new file and then data really disappears.

What is the reason behind this fact? It is just more cost-effective. Overriding the space with zeros would just be one more operation that you would be doing. Hard drives have a limit on writing operations that can be made, so better to override it just once with the new information when need be, instead of doing it twice with the zeros first, and then

[274] Please note that we are using delete and erase as interchangeable expressions. However, in computing terms erasing could mean overwriting the data, while deleting would be just telling the system that there is no information there (as explained below).

with the new information. Furthermore, storage nowadays is so inexpensive that deleting data is not necessarily more expensive than keeping data[275/276].

§ 1.3.3.2. Recovering Deleted Files

Will this operation (deletion) be enough to comply with the right to be forgotten in accordance with the GDPR? Do you need to ensure that deletion is not reversible by advanced tools or (even) theoretical tools? There are good reasons to think that this is not the case. First, to assess whether a person is identifiable one must analyse *"all the means reasonably likely to be used, such as singling out, either by the controller or by another person to identify the natural person directly or indirectly. To ascertain whether means are reasonably likely to be used to identify the natural person, account should be taken of all objective factors, such as the costs of and the amount of time required for identification, taking into consideration the available technology at the time of the processing and technological developments"* (recital 26 of the GDPR). So, in this case, data is definitely still there. Nonetheless, it exists on a "transitional" stage waiting to be deleted. For the data controller recovering the data would mean incurring into high costs, would be time costly and would consume unnecessary computer power (thereby, being environmentally unfriendly). A third-party that has access to the data (through legal or illegal[277] means) would face broadly the same challenges.

The WP29 does not include a controller-centred subjective criteria in its assessment, but we feel that it would be important to do so. When a data controller has a database for its exclusive access, and applies all the technical and organizational measures to ensure data security[278], compliance with the GDPR should not be dependent on the

[275] *See*, Eduard Fosch Villaronga, Peter Kieseberg and Tiffany Li, "*Humans Forget, Machines Remember: Artificial Intelligence and The Right to Be Forgotten*" Computer Law & Security Review 32,2 (2018):304-313; Nadezhda Purtova, *"The law of everything. Broad concept of personal data and future of EU data protection law"* Law, Innovation and Technology 10,1 (2018): 40-81; "Guía de Privacidad desde el Diseño", AEPD, accessed 19 October 2019, https://www.aepd.es/media/guias/guia-privacidad-desde-diseno.pdf.

[276] Even consumer grade hard drives are currently so reliable (the average hard disk will work for 33 years of continuous operation) that Google chose this type of drive instead of more expensive models. In fact, when it preserved about 100.000 terabytes of data, the company did not spend more than a few hundred million dollars in storage. Nowadays, the value of hard drives is considerably cheaper than at the time, with a 1tb hard drive going for values of 40€. *See*, Viktor Mayer-Schönberger, *delete: The Virtue of Forgetting in the Digital Age* (New Jersey: Princeton University Press, 2009): 37-58.

[277] According to the ECJ's judgment in *Breyer* the fact that something is prohibited by law makes it less likely due to '*a disproportionate effort in terms of time, cost and man-power",* making the risk of identification, in fact, insignificant. *See*, Judgment of the ECJ of 19 October 2016, *Patrick Breyer v. Bundesrepublik Deutschland*, Case C-582/14, ECLI:EU:C:2016:779.

[278] *i.e.* the risk of organisational dysfunctions is low.

means reasonably likely to be used to identify a natural person by a third-party, such as a large technological company or an hostile governmental actor[279]. If the database is for its own exclusive access, adequate security measures have been taken and the subject cannot be reasonably identified by the data controller (*i.e.* the controller cannot access it) and it is just waiting to be overwritten, data should, for all intents and purposes be considered to be erased[280].

§ 1.3.3.3. Inputting Data and Deleting Data in AI

Forgetting may have unpredictable consequences. One may not even remember the name of his primary school teacher, but surely does not forget how to read and write because of that. However, if one forgets the rules of basic arithmetic, one will also find itself unable to understand algebra or calculus. This is because knowing the rules of one is essential to understanding the others and without the basis, a gap in knowledge would manifest itself. Information is critical and by not having enough, or by forgetting we can arrive at wrong conclusions (that may not necessarily be illogical). A thought experiment seems in store: A Doctor is examining a patient with a strange condition. Due to some explained reason this patient forgot every blond person that she ever met. Looking at his Doctor (who is blond) the patient immediately asks him if he dyed his hair.

This conclusion seems ridiculous but is far from not logical. From her pool of knowledge, our patient knew that: 1. She knows hundreds of people with natural brunette or red hair; 2. She knowns none with blond hair; 3. It is possible to dye your hair in a variety of colours. In our society, wrong conclusions derived from insufficient or incorrect information are a root of many our problems, such as racism and xenophobia. Ideally, we would like to avoid such pitfalls when designing artificial intelligence.

As we know, ML algorithms base their learning capability on the data that is fed to them. The algorithm uses the data to learn rules that are appropriate to its function. If you

---

[279] In line with these conclusions, the "motivated intruder" test from the ICO only assumes that the motivated intruder is reasonably competent and not an expert hacker. *See*, "Anonymisation: managing data protection risk code of practice", ICO, accessed September 6, 2019, https://ico.org.uk/media/1061/anonymisation-code.pdf.

[280] *See*, "Opinion 05/2014 on Anonymisation Techniques", WP29, accessed September 5, 2019, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf, "Opinion 4/2007 on the concept of personal data ", WP29, accessed September 6, 2019, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf.

desire to make an algorithm learn how to play chess it would be ideal for it to have access to as many played games as you can provide it. Is the model (the result of the training process[281]) personal data? This is the first challenge we need to overcome to "reconcile" AI and the GDPR.

The answer appears to be negative. More, we would be tempted to affirm that the larger the dataset, the lower the likelihood of the information "learned" to be considered itself personal data. According to the GDPR personal data *"means any information relating to an identified or identifiable natural person; an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person"*. Let us imagine that we give an algorithm a million photographs of people so it can learn how to identify when they are smiling and have their eyes open for the camera. We will then implement this information in our camera application for smartphones (let us call it "uSmile"). The model created by our algorithm is a success and it works perfectly. Now one should ask: why are our rules not personal data? Because it is impossible to identify from where our algorithm "learned" how to identify a specific type of smile or eye colour. If we cannot "trace our steps" back to a natural person, the rules are not personal data. In this case, trying to do so would not just be time-consuming and costly. It would be, in all likelihood, impossible. Chances are that even the programmer knows very little about how the AI arrived to that model. That is a problem in itself (see our considerations on black box algorithms above). Problem or not, in this case, it works in AI's favour[282].

In fact, one could argue that feeding the data to a ML algorithm, such as a deep learning algorithm, would be, in fact, be a technique of data anonymisation. Anonymised data is data rendered anonymous "*in such a manner that the data subject is not or no longer identifiable*" (recital 27 GDPR). In analysing data from personal identifiable information from hundreds, thousands or even millions of data subjects, a specific model would be derived that does no longer contains data that that can refer to an identifiable person. This

[281] *See,* note 43.

[282] Surely more could be said about the nature of personal data, but in the end the issue of identifiability is the key to qualify the model derived from the datasets as not personal data and exclude the GDPR's applicability to them. Of course, the data from the datasets has to be properly collected and sanctions can still be applied if it is not. The European legislator clearly intended to avoid the application of the GDPR to non-personal data, even if it was derived from personal data. That is why the GDPR is not applicable to anonymized data.

is the position of Datatilsyne)[283], which considers that models will generally contain an aggregate representation of all data in the system from where it would not be possible to retrieve personal data[284]. The characterisation of the technique as aggregation seems doubtful to us, as aggregation generally implies that the aggregated data is made available in a single place/database and relatively reachable. In the case of ML, the representation of the aggregated data is spread somewhere within the resulting model and in a manner which may not even be understandable. Thus, the anonymisation of rules derived from the algorithm would probably be much more difficult (if not impossible) than any normal aggregation.

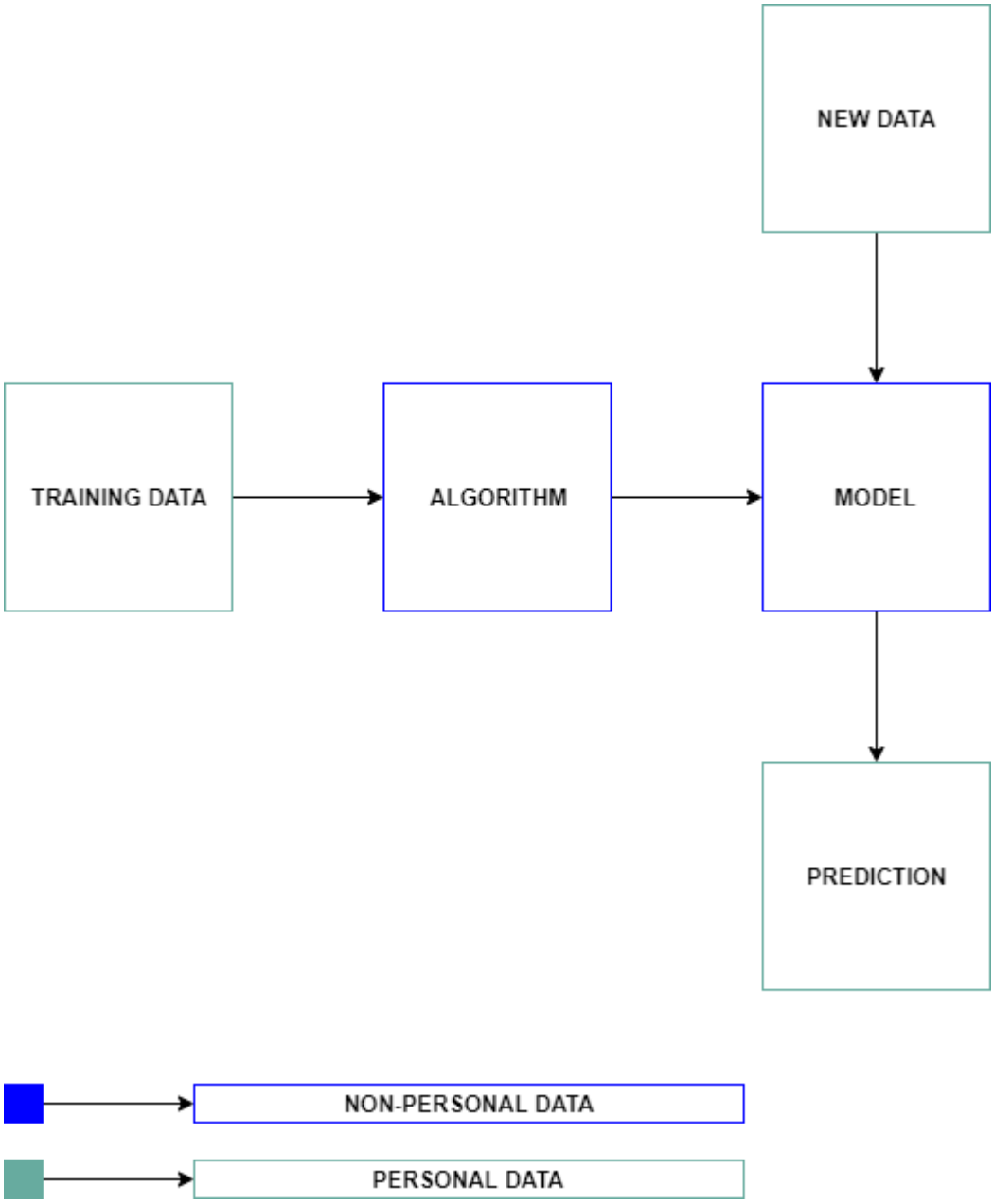
Obviously, this rational is not necessarily applicable to the results (the conclusions/the inferences) arising from the application of the model to new data. If a banking institution uses an algorithm to generate a model for attributing credit ratings to new clients requesting loans connecting the client with a rating. Said credit rating is undeniably personal data. There is personal data in datasets used to train the algorithm. However, and opposite, it is unlikely that there is personal in a model. Nevertheless, personal data will be present again in the credit scores given to the clients (the information inferred by applying the model). If some of the requisites for data erasure is applicable to them, data must still be erased (without prejudice to the right to object and to not be subjected to decisions based on automated decision making when applicable)[285]. If the model is producing incorrect answers data subjects will still have the right to not be negatively affected by the results and to avoid costs with excessive requests the data controller should endeavour to fix it.

---

283 "Artificial intelligence and privacy", Datatilsynet, accessed May 15, 2019, https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf.

284 According to Datatilsynet, *"Decision trees represent one exception to this, as they contain a varying degree of the model's data basis. The limits here depend on whether the tree is "pruned" after learning, or a level limitation is set for learning. One or the other will normally be chosen, as the model should generalise and not over"*.

285 Even though some Authors express doubts on the applicability of personal data to inferences, it does seem to be clear that is applicable when there is personal data in said inferences. The fact, that the right to data portability does not include the data that is created by the data controller according to the WP should not be interpreted as having any relevance for the case of erasure or rectification. One should not forget that data portability is restricted by design to data that is provided by the data subject. In its guidelines on data portability the WP29 is considering only the specific issue of what "data provided by the data subject" means in the context of the GDPR. *See*, Lilian Edwards and Michael Veale, *"Slave to the Algorithm? Why a 'Right to an Explanation' is Probably not the Remedy You are Looking For"*, Duke Law & Technology Review 16,1 (2017): 18-84; "Guidelines on the right to data portability", WP29, accessed May 20, 2019, https://ec.europa.eu/newsroom/document.cfm?doc_id=44099,



In the diagram above, you may find a graphical representation of what constitutes personal data and, thus, falls under the GDPR's (including the right to erasure) rules and what does not. This is, of course, a very general example, and analysis would depend on a number of facts such as the presence of personal data in the datasets, type of algorithm etc., but should hold for most situations[286]. Obviously, it is also possible to have inferences

[286] In its report on AI and Privacy, Datatilsynet has some interesting graphical representations of machine learning, in which we based this example and added the indication of where there is personal data. Thus, we highly recommend consultation to the abovementioned document.

that do not contain personal data. Imagine that a political campaign is running an analysis of voter's opinion on a determined subject by mining data from social networks. The algorithm knows that the use of certain expressions and sentences means that someone supports position A (cats are amazing) or position B (cats are only pretty good). If it is designed to have as an output just the aggregated data of all the opinions it analyses (e.g. 55% in favour, 30% against and 15% do not know/do not answer), said inferences would not include personal data and thereby would need to be erased.

Since anonymisation is, in itself, a data type of data processing one would need a legal basis to anonymise the data in this manner. Legitimate interests of the data controller should suffice in most cases, and when not applicable consent of the data subject could be relied upon, since any withdrawal of consent would not affect the rules inferred (or even the performance a contract).

With the first problem (the easy one) behind us, we should not proceed to the second. Even if the mathematical models derived from the datasets themselves are relatively "safe" from erasure, the same cannot be said about the datasets themselves. Those can, indeed, contain personal data and if there is a request for erasure that meets the legal requirements data controllers will have to comply with it. How could that create a problem? That is relatively simple, to achieve the best results the most varied datasets are needed and the datasets and algorithms themselves may need to be fine-tuned several times to achieve optimum results. While it is not certain that removing just a single input (or even several) from the dataset will significantly hinder the development, if enough data subjects with the same characteristics request erasure the results and rules derived may differ. If the datasets are constantly altered, it may not be possible to perfect them and the algorithms/models.

Imagine an evolutionary algorithm, this algorithm created the best mathematical model at identifying cats in a dataset X (by 1%). However, you suddenly remove 50 of the 100 pictures of cats of that dataset and replace them with 50 new pictures[287]. There will be some entropy, maybe the algorithm is not as good with those new pictures, maybe one of the algorithms who "lost" the evolutionary race would, in fact, be better with this new dataset.

With an algorithm based on analogy, the results could, also, be quite catastrophic. You are using your algorithm to learn to identify faces in photographs. Unfortunately, you

[287] Pictures of cats are not personal data, but it is a good example, nonetheless.

were not able to get a large sample of a certain minority group, but its size is large enough and luckily your algorithm is effective so, for now, it works for said minority. Still, you need to tweak your algorithm to correct some imperfections and you plan on feeding it the same dataset in the future. However, while you do that a certain number of the data subjects from that minority ask you to erase their data, and you comply. With this difference in data, your improved algorithm and new model is suddenly unable to identify people from that minority group. You manage to overcome whatever imperfections it had and, technically, it works, but due to the diminished size of your sample it now discriminates against a certain group. It is possible to find new subjects, of course, but that will cost money and delay the development process. When development is a constant procedure and requests for erasure are also frequent it could create serious entropy or render the whole process unfeasible.

Lastly, you are teaching a deep learning algorithm to do automatic machine translation. To do so, you ask the users of your website to upload their previous translations of documents that they have made and put all this documentation on an ever-growing database against which the algorithm is constantly checking and updating itself. When there is personal data in the documents, you ask for consent as your legal basis for the processing. Originally, you have a great number of PT to ENG submissions and, thus, your algorithm is relatively proficient ins translating from these languages. However, after a few months, your users start withdrawing consent and you lose a significant number of your data on PT to ENG translation. What would be the consequences? Best case scenario the rules learned by the algorithm are enough so it will not regress but is likely that any progress will take some time and only happen with a new dataset. Worst case scenario it completely breaks your algorithm since it is constantly checking against and updating itself using that database. Obviously, the consequences would depend on the specific case, and we are certainly not equipped to analyse them. But, in abstract, all these are possible consequences.

The issue will be more challenging to overcome in non-static algorithmically created ML-models. That is ones that are constantly changing in accordance to new data provided. If they are static, worst case scenario, when they are replaced by a new mathematical model (the only option that allows for change) the new one is not as accurate as the previous. If they are non-static, it could completely break the model.

Then, how to avoid disruptions of databases and a potential chilling effect on AI caused by the right to erasure? For data processors the advice is to take into account the principles of privacy by design and privacy by default and try to build algorithms that are resistant to the erasure of certain data entries. Making sure you keep large and varied samples may also help to minimise the damage caused by the erasure of some data. Taking into account the principle of data minimisation, when it is possible to work absent personal data or with anonymised data that step should be taken. When it is not possible trying to use as little personal data as possible is ideal[288]. Finally, when possible, avoid consent and legitimate interests as your legal basis. For example, if you are working on a self-driving vehicle, you can opt to get your sample by contracting drivers to drive cars and get data for you. In that manner, the legal basis will be the performance of the contract where the right to erasure is more restricted[289].

Courts and data protection authorities should take a realistic approach to the issue and, minding arguments above, consider that the GDPR is not applicable to the models derived from the original personal data. Furthermore, the expression "undue delay" could be interpreted in a manner that considers the time needed for the data controller to adapt its technologies to the erasure of the personal data. That time could vary in accordance

[288] Obviously, the "less is more approach" will not always work for AI and there will be situations when "more is more".

[289] *See,* "The Impact of the EU's New Data Protection Regulation on AI", Nick Wallace and Daniel Castro, accessed August 9, 2019, http://www2.datainnovation.org/2018-impact-gdpr-ai.pdf; "The Next Big Privacy Hurdle: Teaching AI to Forget", Darren Shou, accessed September 7, 2019, https://www.wired.com/story/the-next-big-privacy-hurdle-teaching-ai-to-forget/; Matthew Humerick, *"Taking AI Personally: How the E.U. Must Learn to Balance the Interests of Personal Data Privacy & Artificial Intelligence"* Santa Clara High Technology Law Journal 34,4 (2018): 393-418. Regarding this last paper, its analysis of the conflicts between the rights enshrined in the GDPR and AI is interesting, even if the Author shows a clear lack of knowledge about EU trade policy and the single market. For clarification purposes, while the UK is, indeed, a very important player on AI it is very likely that being part of the single market plays a central role in it. If the UK leaves the EU with a deal that allows for a high degree of access to the single market it is highly likely to include maintaining similar standards for fundamental rights, including data protection. Even without a deal, if the UK wishes to have an adequacy decision from the European Commission (without which its AI companies would be in serious jeopardy) it will need to have similar standards to the EU. Therefore, relaxing them [the standards] too much is not really an option. There is also the fact that the UK is a country that is sensible to data protection concerns, its supervisory authority (ICO) is one of the best equipped in the entire European Union. In fact, as of this writing the two highest fines under the GDPR were levied by the ICO: £183.39 million against British Airways and £99 against the hotel giant Marriott. Implying that data protection standards in the UK would lower with Brexit is, therefore, no more than guesswork and quite poorly informed guesswork at that. Besides, affirming that law-making in any Member State is a more streamline procedure than at the European level is generally, but not necessarily true. Law-making at the EU level faces the usual challenges existent in any federal or quasi-federal entity. On the other hand, EU law in core issues is relatively stable and not as much at the mercy of ephemeral majorities. Last, but not least, we cannot agree with quoting studies about the impact pre-brexit of AI's impact on the economy as if the fact that there may be a Brexit does not affect the results of those studies (at least a disclaimer would advisable).

with the technology used and would probably need to be evaluated on a case-by-case basis[290].

§ 1.4. The Right to Rectification

Article of 16 of the GDPR, bestows upon the data subject the right to obtain rectification of data concerning him/her and to have incomplete personal data completed. The issues around the right to rectification are similar to the ones arising from the right to erasure, including, but not limited to the question of the nature of the rules derived from algorithmic processing. As with the right for erasure, if the model does not contain personal data, the data subject does not appear to have any right to ask for rectification or completion of the data contained in said model. Still, the right subsists for the original dataset, with the limitations abovementioned. There are three more aspects that should be taken into account: on one hand, even if there is no GDPR-mandated obligation to rectify there is certainly an economic incentive to do so. After all, an algorithm that works with incorrect data will (probably) produce inferior results.

The legal basis for data processing may also be tainted by the fact that data is not correct. Of course, this is dependent on a casuistic approach, and the collecting of the incorrect data by the data controller probably has to be, at least, negligent to justify any practical consequences from this fact. Nonetheless, it is undeniably that the legitimate interest of the data controller in collecting imprecise data is inferior to the interest in collecting correct data. That may be enough to tip the balancing test[291]. The same could be said about the performance a contract. Arguably, imprecise data is not needed to perform a contract and it actually may hinder said performance.

As a follow-up to the last point, it is recommendable that data controllers in the area of AI development do not forget that the principles of accuracy and data minimisation are general principles of the GDPR. Infringement may result in fines of up to 20€ Million

[290] Law-making suggestions shall be left to the conclusions.
[291] "Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC", WP29, accessed June 8, 2019, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf; Information Commissioner's Office, *Guide to the General…*: 81ff; "Processing Personal Data on the Basis of Legitimate Interests under the GDPR", Gabriela Zanfir-Fortuna and Teresa Troester-Falk, accessed June 9, 2019, https://info.nymity.com/hubfs/Landing%20Pages/Nymity%20FPF%20-%20Legitimate%20Interests%20Report/Deciphering_Legitimate_Interests_Under_the_GDPR.pdf?hsCtaTracking=9cf491f2-3ced-4f9c-9ffa-5d73a77a773e%7C7469b2ec-e91c-4887-b5db-68d407654e23

or 4% of the total worldwide annual turnover of the preceding financial year, whichever is higher.

§ 1.5. The Right to Object

Data subjects have the right to object to data processing activities (including profiling) based on legitimate interests of the data controller (or a third party) or in the necessity for the performance of a task carried in public interest or in the exercise of official authority vested in the data controller. When such a situation arises, the data controller must stop all processing activities unless it can demonstrate legitimate grounds that override the interests, rights and freedoms of the data subject (or when the processing is necessary for the establishment, exercise or defence of legal claims).

Another difficult balancing test will be in store when a data subject exercises their right to object. It is not an absolute right, but the arguments stated above must be taken into consideration when deciding if the requirement must be fulfilled. Still, data controllers using ML algorithms would be well advised to try to develop privacy-compliant AI solutions, because refusal to carry out a request on the basis of costs will probably become less acceptable as time goes on and technology progresses.

One relevant exception is that the right to object is absolute when data is used for direct marketing purposes. If applicable, the controller has no grounds to object and must cease the processing activities, taking also into account Article 17(1), point c) of the GDPR. Since the information inferred by applying the model will probably have personal data when processing is for direct marketing purposes, processing of this information has to cease. Exercise of this right by the data subject only affects future processing and has no bearing on past processing activities[292/293].

[292] Our considerations about the model and dataset in the section regarding the right to erasure are still applicable.
[293] Information Commissioner's Office, *Guide to the General…*: 144ff.; Paul Voigt and Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR)…* :176-181.

§ 1.6. The Right to Explanation

§ 1.6.1. Legal Basis under the GDPR

We may find relevant references to the right to explanation in three different sets of provisions of the GDPR: *a)* the right to information (Articles 13 and 14 of the GDPR); the right to access (Article 15 of the GDPR) and; the right not be subject to automated individual decision-making, including profiling (Article 22 of the GDPR).

As an introduction, it is necessary to go into the different provisions and see how they might differ.

In accordance with the mandatory information to be provided to data subjects under the GDPR, the data subject has the right be informed of *"the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4)*[294] *and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject"*.

The right to access states that *"the data subject shall have the right to obtain from the controller confirmation as to whether or not personal data concerning him or her are being processed, and, where that is the case, access to the personal data and the following information"* including *"the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject""*. On this matter, Article 15 is further complemented by Recital 63 according to whom *"every data subject should therefore have the right to know and obtain communication in particular with regard to the purposes for which the personal data are processed, where possible the period for which the personal data are processed, the recipients of the personal data, the logic involved in any automatic personal data processing and, at least when based on profiling, the consequences of such processing."*

In accordance with Article 22 of the GDPR (the right not to be subject to automated individual decision-making, including profiling), data subjects have the right not to be subjected to a decision based solely on automated processing which produces legal effects concerning him or her or similarly significantly affects him or her. There are three exceptions to this rule: *i)* explicit consent; *ii)* when it is necessary for entering into or performing a contract, where the controller and data subject are parties *or; iii)* is authorised

[294] General prohibition on certain types of automated decision-making.

by EU or Member State law. In every case suitable safeguards must be implemented and, at least when in the context of consent or contract the data subject must be able to obtain human intervention. Further restrictions are in place for special categories of personal data, in this case it is only admissible through specific consent or reasons of substantial public interest (and, of course, pursuant to abovementioned suitable safeguards). Recital 71 complements this provision with the following *"such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision"*

§ 1.6.2. Preliminary Question: What is a Based Solely on Automated Processing?

Some controversy arose about what should be interpreted by "based solely on automated decision". In trilogue negotiations the EP proposed a broader formulation which included "solely or predominantly", and the apparently restrictive final choice was the source of some degree of confusion. Did the European legislator intent to enshrine a regime where the mere fact that a human participated in the decision-making excludes the application of Article 22? Even if said human was only rubber-stamping the conclusions of the machine?

From the beginning, that interpretation did not seem acceptable. The use of the adverb solely implies that a decision does not involve anyone or anything else. That is to say it is exclusively derived from the machine's decision. Predominantly, on the other hand, would make the Article be applicable to any decision where 51% of the final result was due to the machine's decision. That is to say, using the word predominantly could give rise to undesirable situations. Imagine that there is a human supervising the algorithm, said human has perfect knowledge of the rules to be applied and of the algorithm's working and possesses the power to change its decision if he/she decides to do so. For argument's sake, said human actor is working diligently in supervising the algorithm. Under the predominantly formulation, Article 22 would still be applicable, since most of the work would still be done by the algorithm (it is unlikely that the human can review by doing the same amount of calculations, the human actor will probably try to find mistakes in the algorithm's inner workings and look for nuances that the algorithm did not understand). The result of such an option could be that a much higher number of data processing

activities would be restricted under Article 22, causing significant problems for data controllers, potentially even regarding legal basis for data processing activities[295]. Further, it would probably remove all incentive to have a human actor "controlling" algorithmic decisions and may even be counterproductive. In fact, if Article 22 is always applicable data controllers could think that might as well cut costs and go for entirely automated means.

As it is possible to conclude, quite a few problems could appear by using the word predominantly and those, by themselves, justify its suppression. It does not mean that the European legislator intended to apply an ultra-restrictive approach to human intervention. In fact, if the legislator intended to do so, there are other redactions that would be far more effective[296] or, in alternative, the general prohibition could be suppressed since it would lose all its effectiveness. Put a person "approving" a million algorithmic decisions in bulk and you would have the problem solved. Evidently, we should not forget that such an interpretation would go against the clear rationale of the GDPR.

Solely implies, realistically, that the power to make the decision rests exclusively in the program. If the person who is supervising it does not have the power to override or does not understand its functioning in a manner that would allow him/her to detect errors[297] and correct them, said power still sits solely with the program and, thereby, Article 22 is applicable. The analysis should be, as with the establishment of the legal position of the parties as data processor or controller, based on the situation *de facto* and not *de jure*.

In its Guidelines on Automated individual decision-making and Profiling, the WP29 clarified its opinion about what it understood by human involvement. In line with the above explained, it considers that data controllers cannot fabricate human involvement to avoid the application of Article 22. According to the Guidelines *"to qualify as human involvement, the controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision. As part of the analysis, they should consider all the relevant data"*. While decisions from the WP29 are not binding, they represent the shared opinion of data protection supervisory authorities across the European Union and we find highly unlikely that, if called upon to

[295] You would never be able to rely on the legitimate interests of the data controller for any processing activity where an algorithm was used for a pre-assessment, even if a human had key intervention at a second stage.
[296] Something along the line of "data subjects have the right not to be subjected to a decision based on automated processing, save if any type of human overlook is established".
[297] In fact, the person should able to detect errors, at least, at the level of someone who is reasonably familiarized with the program and is reasonably diligent.

do so, the European Court of Justice, taking into account its tradition in protecting data protection and privacy rights, would opt for a different interpretation (and less strict). The controversy seems to be, by now, put to rest[298].

§ 1.6.3. Preliminary Question: What is a Decision that Produces Legal Effects Concerning him or her or Similarly Significantly Affects him or her

Likewise, the GDPR offers no definition about what should be considered as a decision producing legal or similarly significant effects. Recital 71 offers the example of *"automatic refusal of an online credit application or e-recruiting practices without any human intervention"* but it is not possible to find further clarification of the legal text.

Again, the input of the WP29 is invaluable to provide some degree of clarity. In accordance with its Guidelines on automated individual decision-making, a decision shall be considered has having legal effects when it affects a person's legal rights, legal status or rights under a contract. The ICO adds that to trigger the application of Article 22 and the general prohibition said legal effects have to be adverse to the data subject. However, and as reasonable as this seems at first glance, one has to note that it can be rather troublesome. What is an adverse effect on the legal rights of the data subject? If there are three job openings for a position at a company, and it uses an algorithm to select candidates, the one that managed to get third place did not suffer an adverse effect *stricto sensu*, but there may some interest in, for example, obtaining human intervention if the candidate feels like he/she should have been seriated first.

The definition of an effect that is similar is slightly more difficult and must be accessed on a case-by-case basis. According to the Guidelines on automated individual decision-making, the threshold should be that those effects should affect the data subject in a similar manner as the legal effects. The WP29 establishes three criteria: the decision must have the potential to *i) "significantly affects the circumstances, behaviour or choices of the individuals concerned"; ii) "have a prolonged or permanent impact on the data subject" or; iii) "at its most*

[298]*See,* Paul Voigt and Axel von dem Bussche, *The EU General Data Protection Regulation*…, 181ff.…; "Opinion 1/2010 on the concepts of "controller" and "processor"", WP29, accessed May 5, 2019, https://ec.europa.eu/justice/article-29/documentation/opinion recommendation/files/2010/wp169_en.pdf; "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", WP29, accessed May 5, 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826

extreme, lead to the exclusion or discrimination of individuals"[299]. The similarly significant effect can arise from situations created by third parties and not necessarily by the data subject. The WP29 gives the example of credit card company reducing a customer's credit card limit based on the financial history of people in the same residential area.

While a considerable amount of discussion is still possible regarding what should be considered as having a similarly significant effect on specific cases it seems extremely unlikely that this (highly reasonable) interpretation will be struck down by ECJ in favour of one that is less protective of data subjects[300]. Therefore, following it seems the most reasonable choice[301/302].

§ 1.6.4. Preliminary Question: Recitals Under EU Law

Article 296 TFUE states that EU acts *"shall state the reasons on which they are based and shall refer to any proposals, initiatives, recommendations, requests or opinions required by the Treaties"*. Thereby, recitals (and citations) are mandatory for EU legal instruments and in their absence, the Act is void. The requirement is applicable to every act from EU institutions, bodies, offices or agencies that are intended to produce legal effects in relation to third parties. The European Court of Justice has previously stated that the *raison d'être* of this disposition is to allow the Court itself its power of review and Member States and parties

---

[299] The WP29 gives a few examples of decisions that should be considering as meeting this threshold such as:

a) *"decisions that affect someone's financial circumstances, such as their eligibility to credit;*
b) *decisions that affect someone's access to health services;*
c) *decisions that deny someone an employment opportunity or put them at a serious disadvantage;*
d) *decisions that affect someone's access to education, for example university admissions".*

[300] While the ECJ does not defer to the EDPB on matters of data protection, its undeniable that the Board's opinions have persuasive authority over both the Court and the opinions of the Advocate Generals' frequently followed by the Court. We provide below a few examples where the Guidelines or Opinions were cited by Court or (more frequently) Advocate Generals. Further, as far we know there is no single instance where the ECJ contradicted the EDPB or its previous itineration the WP29, to establish a rule that protected data subjects in a minor degree. *See*, Judgment of the ECJ of 10 July 2019, *Tietosuojavaltuutettu v. Jehovan todistajat — uskonnollinen yhdyskunta*, Case C–15/17, ECLI:EU:C:2018:551; Judgment of the ECJ of 30 May, 2006, *Parliament v. Council*, Joined Cases C-317/04 and C-318/04, ECLI:EU:C:2006:346. As abovementioned, the situation is more frequent for AG opinions, with 16 references. *See*, Rita de Sousa Costa, A realização do direito da protecção de dados da União Europeia através das fontes não-legislativas: dos grandes temas jurisprudenciais aos desafios do soft law, no contexto da aplicação do Regulamento Geral sobre a Protecção de Dados (Master's thesis: Universidade Católica Portuguesa, 2019).

[301] Plus, no reasonable entity should implement compliant measures against the collective interpretation of the data protection supervisory authorities

[302] *See*, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", WP29, accessed May 5, 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826; Information Commissioner's Office, *Guide to the General …,* 155.

concerned to understand the manner in which the Institutions have applied the Treaties (the central question seems to be compliance with the principle of subsidiarity)[303].

Now that the reason behind the existence of recitals is explained, what is their value in interpretation? It is widely accepted by the ECJ's case law that recitals are not binding. In fact, the Court states that *"the preamble to a Community act has no binding legal force and cannot be relied on as a ground for derogating from the actual provisions of the act in question"*[304].

Recitals cannot derogate operative provisions. They can, however, provide clarification on certain aspects of those provisions, including their scope or purposes[305]. Thus, if a recital contains a rule that is not reflected in any of legal instrument's enacting terms (the substantive provisions or articles) said rule will not have legal binding value. If a recital contains a provision that is clearly contradictory with the articles, the latter shall prevail. However, it is incorrect to think that recitals have no value. They may provide clarification *"on the interpretation to be given to a legal rule, [even if by themselves they] cannot constitute*

---

303 *See*, Tadas Klimas and Jūratė Vaičiukaitė, *"The Law of Recitals in European Community Legislation"* ILSA Journal of International & Comparative Law 15,1 (2008): 1-31; Ingrid Opdebeek and Stéphanie De Somer, *"The Duty to Give Reasons in the European Legal Area a Mechanism for Transparent and Accountable Administrative Decision-Making? - A Comparison of Belgian, Dutch, French and EU Administrative Law"* Rocznik Administracji Publicznej 2 (2016): 97-148; Manuel Lopes Aleixo, "Anotação ao art.º 296.º do TFUE", in *Tratado de Lisboa*…, 1060-1062.
The ECJ had long recognized that the reasoning behind Acts is a mandatory formality, including the statement of the legal basis pursuant to which the act is adopted. *See*, Judgment of the ECJ of 8 February 1968, *Fonderie Acciaierie Giovanni Mandelli v. Commission of the European Communities*, Case 3/67, ECLI:EU:C:1968:6; Judgment of the ECJ of 16 June 1993, *France v. Commission*, Case C-325/91, ECLI:EU:C:1993:245.

304 *See*,Judgment of the ECJ of 19 November 1998, *Nilsson*, Case 162/97, ECLI:EU:C:1998:554.

305 *See*, Judgment of the ECJ of 27 November 2007, *C.*, C-435/06, ECLI:EU:C:2007:714:
51. *"The term civil matters' must be interpreted as capable of extending to measures which, from the point of view of the legal system of a Member State, fall under public law"*
"That interpretation is, moreover, supported by Recital 10 in the preamble to Regulation No 2201/2003, according to which that regulation is not intended to apply 'to matters relating to social security, public measures of a general nature in matters of education or health ...' Those exceptions confirm that the Community legislature did not intend to exclude all measures falling under public law from the scope of the regulation"
Judgment of the ECJ of 26 June 2001, *The Queen v Secretary of State for Trade and Industry, ex parte Broadcasting, Entertainment, Cinematographic and Theatre Union*, Case C-173/99, ECLI:EU:C:2001:356.
"37. As regards, first, the purpose of Directive 93/104, it is clear both from Article 118a of the Treaty, which is its legal basis, and from the first, fourth, seventh and eighth recitals in its preamble as well as the wording of Article 1(1) itself, that its purpose is to lay down minimum requirements intended to improve the living and working conditions of workers through approximation of national provisions concerning, in particular, the duration of working time.
38. According to those same provisions, such harmonisation at Community level in relation to the organisation of working time is intended to guarantee better protection of the health and safety of workers by ensuring that they are entitled to minimum rest periods and adequate breaks.
39. In that context, the fourth recital in the preamble to the directive refers to the Community Charter of the Fundamental Social Rights of Workers adopted at the meeting of the European Council held at Strasbourg on 9 December 1989 which declared, in point 8 and the first subparagraph of point 19, that every worker in the European Community must enjoy satisfactory health and safety conditions in his working environment and that he is entitled, in particular, to paid annual leave, the duration of which must be progressively harmonised in accordance with national practices"
Gianclaudio Malgieri and Giovanni Comande, *"Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation"* International Data Privacy Law 7,4 (2017): 243-265

such a rule" as was established by the ECJ in its *Casa Fleischhandels* judgment. The Court was careful in stating that a *"recital cannot be relied upon to interpret [an Act] in a manner clearly contrary to its wording"*. Thus, *a contrario*, a recital may be relied upon to interpret Act in a manner that is not contradictory to its wording, giving them explanatory value. Thereby, it seems fair to argue, from the ECJ's case-law, that while not binding, recitals are a powerful tool for interpretation. Indeed, the recitals are frequently a direct line to the intentions behind the norm"[306].

Some Authors even point that they may play the role of supplementary normative tools, basing such argument, in part in the Commission's Communication on guidance for better transposition and application of Directive 2004/38/EC. We cannot find a legal foundation neither in the Union's Constitutional law nor in the ECJ's case-law to support

[306] *See*, "Complexity of EU law in the domestic implementing process", Roberto Baratta, accessed May 2, 2019, http://ec.europa.eu/dgs/legal_service/seminars/20140703_baratta_speech.pdf; Roberto Baratta, *"Complexity of EU Law in the Domestic Implementing Process"* The Theory and Practice of Legislation 2,3 (2014): 293-308; Gianclaudio Malgieri and Giovanni Comande, *"Why a Right to*…, 243-265; Judgment of the ECJ of 25 November 1998, *Giuseppe Manfredi v. Regione Puglia*, Case C–308/97, ECLI:EU:C:1998:566; Judgment of the ECJ of 13 July 1989, Casa Fleischhandels v. Bundesanstalt für landwirtschaftliche Marktordnung, Case 215/88, ECLI:EU:C:1989:331.

this position. In this specific case, the Commission's is far from clear (and stable) and, even if it was, its interpretation holds no special value in this matter[307/308].

§ 1.6.5. Preliminary Question: Algorithmic Bias and why the Right to Explanation is a Key Tool to Fight it – 4 Stories

§ 1.6.5.1. The Misogynistic AI

Hiring the employees that best fit your company's necessities is quite challenging and taxing work. If the position is desirable enough, recruiters may have to sort out through thousands of CVs, motivation letters and other documentation from candidates. And that is just the first stage, the road until you are talking to face-to-face with your potential new colleague is a long one. Thereby, it comes as no surprise that HR departments would be thrilled to have a tool that would ease somewhat this burden and

---

[307] *See*, Tadas Klimas and Jūratė Vaičiukaitė, *"The Law of Recitals*…, 1-31; Gianclaudio Malgieri and Giovanni Comande, *"Why a Right to Legibilit*…: 243-265; Llio Humphreys et al., "Mapping Recitals to Normative Provisions in EU Legislation to Assist Legal Interpretation", *in Legal Knowledge and Information Systems* Vol. 279 Antonino Rotolo ed., (Amsterdam: IOS Press, 2015), 41-49; Sandra Wachter, Brent Mittelstadt and Luciano Floridi, *"Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation"* International Data Privacy Law 7,2 (2017): 79-99; "Introduction to the legislative processes for EU directives matters and regulations on financial services", Slaughter and May, accessed May 4, 2019, https://www.slaughterandmay.com/what-we-do/publications-and-seminars/publications/client-publications-and-articles/i/introduction-to-the-legislative-processes-for-eu-directives-and-regulations-on-financial-services-matters.aspx; Wim Voermans, Maarten Stremler and Paul Cliteur, *Constitutional Preambles: A Comparative Analysis* (Northampton: Edward Elgar Publishing, 2017): 11ff. "Communication from the Commission to the European Parliament and the Council on guidance for better transposition and application of Directive 2004/38/EC on the right of citizens of the Union and their family members to move and reside freely within the territory of the Member States", European Commission, accessed May 3, 2019, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52009DC0313&from=EN; Judgment of the ECJ of 1 October 2009, *Commission v. Council*, Case 370/07, ECLI:EU:C:2009:590;; Judgment of the ECJ of 21 December 2016, *Club Hotel Loutraki v. Commission*, C-131/15 P, ECLI:EU:C:2016:989; Judgment of the ECJ of 16 June 1993, *France v. Commission*, C-325/91, ECLI:EU:C:1993:245; Judgment of the ECJ of 16 June 1993, *France v. Commission*, C-325/91, ECLI:EU:C:1993:245; Judgment of the EGC of 30 September 2015, *Alexios Anagnostakis v. Commission*, Case T-450/12, ECLI:EU:T:2015:739; Judgment of the ECJ of 12 September 2017, *Alexios Anagnostakis v. Commission,* C-589/15 P, ECLI:EU:C:2017:663; Judgment of the ECJ of 4 July 1963, *Germany v. Commission,* 24/15, ECLI:EU:C:1963:14;, ECLI:EU:C:1963:14; Judgment of the ECJ of 16 June 1993, *France v. Commission,* C-325/91, ECLI:EU:C:1993:245; Judgment of the Court of First Instance of 13 September 1995, *TWD Textilwerke Deggendorf GmbH v Commission,* Joined Cases T-244/93 and T-486/93, ECLI:EU:T:1995:160; Judgment of the ECJ of 15 May 1997, *TWD Textilwerke Deggendorf GmbH v Commission,* Case C-355/95 P, ECLI:EU:C:1997:241; Judgment of the ECJ of 24 November 2005, *Deutsches Milch-Kontor GmbH v Hauptzollamt Hamburg-Jonas*, Case C-136/04, ECLI:EU:C:2005:716.

[308] While we admit that it is enticing, absent a proper sustaining argumentation calling upon a specific Commission's Communication is not persuasive at all. While having an Institution defending this interpretation could be important, mainly because it was very competent professionals working for it and could successfully argue it in the ECJ, as far as we know the abovementioned Communication is an isolated case and it does not seem inclined to defend this position in a consistent manner.

allow professionals to focus on other related tasks, like giving them the opportunity to make a deeper evaluation of (just) the top candidates. Amazon, a leader in AI development, clearly was in a prime position to successfully implement a tool that would succeed in this endeavour. Starting in 2014 it tried to. It failed, badly!

There was one main problem with Amazon's tool: it really did not want to hire women. To develop the tool, data from past hirings was fed to the learner algorithm. Due to the abovementioned problems regarding (the lack of) gender balance in the ICT sector, most of the CVs belonged to male candidates. Some skills that were common across applicants were considered of low relevance, even if they were key for a position in the sector, like the ability to write computer code. Being a woman, on the other hand, was considered highly relevant, in fact, it was weighted as exclusion criteria. The algorithm looked at the low number of women previously hired and (being unable to understand the nuances traditional of gender imbalance in ICT) concluded that that was the result of women not being a good fit for Amazon.

Though one has to admit that the AI was more effective than any human recruiter could ever be, that is at discriminating against women. The program went has far as to detect typical masculine and feminine language and exclude candidates who employed the latter in their CVs (even if they were actively trying to, human recruiters would have a hard time achieving this). When the company introduced safeguards to curb its misogynist tendencies, it stopped working altogether, recommending unqualified candidates for random positions. With unsatisfactory results and no guarantee that, even if Amazon managed to get it on working condition again, the algorithm would not find a manner to circumvent the safeguards and revert back to its old habits, the project was halted.

According to Amazon, the technology never passed the pilot phase, so no real-world damage appears to have materialised, and its failure is a valuable lesson on how AI may reflect the bias contained in the data that is fed to it. The company did manage to salvage some of its progress, using it for less complex tasks such as sorting through duplicate CVs. Amazon is currently trying again at ML-enabled recruiting with a focus on diversity[309].

[309] *See*, "Amazon scraps secret AI recruiting tool that showed bias against women", Jeffrey Dastin, accessed June 11, 2019, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G; "Amazon's Secret AI Hiring Tool Reportedly 'Penalized' Resumes With the Word 'Women's'", Rhett Jones, accessed June 11, 2019, https://gizmodo.com/amazons-secret-ai-hiring-tool-reportedly-penalized-resu-1829649346; "Amazon offers cautionary tale of AI-assisted hiring", Andrew Hill, accessed June 11, 2019,

§ 1.6.5.2. Anti-Semitic AI

In 2016 Microsoft unveiled the chatbot Tay, described in her bio has a "AI fam from the internet that's got zero chill". Tay was supposed to mimic a real human teenage girl and interact with Twitter users in real-time. For what is worth, Tay was somewhat believable in her initial interactions with users, having almost human-like conversations.

In 24 hours, Tay went from being very excited to meet the "super cool" humans to posting anti-semitic, politically inflamed, racist and xenophobic tirades[310]. There were two major reasons for this. First, Tay was designed to learn from her interaction with humans, but not all humans are righteous upstanding citizens (on Twitter of all possible places), and if we program AI's to closely mirror the users' behaviour, some of humanity's worst characteristics are bound to show up. Proving once and for all and beyond a shadow of a doubt that none of us is as bad as all of us, a group of users target Tay with messages with the abovementioned characteristics. Since her filters were not advanced enough to prevent (we will get to that), Tay developed the same "ideology" and started sharing unappropriated messages. The second problem, and a major oversight by Microsoft was that Tay included a feature where it would repeat a message tweeted at her by the user. In hindsight, it is easy to see why that was a really bad idea, and it indeed, worked out very poorly.

Microsoft deactivated Tay once and tried to introduce some safeguards but it was not enough. A second run produced similar results and the company pulled the plug on the bot indefinitely.

Microsoft owned to Tay's faults and seems to be making an effort to rid their AI programs of any undesirable characteristics with some degree of success. However, that success also shows some of the limitations of filtering AI by hand. AI may encounter millions of new situations and a few thousand filters are not enough to cover them all. If they are too broad, they will probably diminish its potential. Case in point Zo, Tay's

https://www.ft.com/content/5039715c-14f9-11e9-a168-d45595ad076d; "Amazon scrapped 'sexist AI' tool", BBC News, accessed June 11, 2019, https://www.bbc.com/news/technology-45809919; "Amazon reportedly scraps internal AI recruiting tool that was biased against women", James Vincent, accessed June 11, 2019, https://www.theverge.com/2018/10/10/17958784/ai-recruiting-tool-bias-amazon-report.

[310] It is our opinion that reproducing the content of the messages posted by the chatbot will not lend any further strength to our arguments and may do no more than causing some discomfort to a section of potential readers. Therefore, we chose not to do it. Still, if you wish to find the content of said messages, and it is important that it is available to fully understand what happened, you may do so in the references provided in footnote 310.

successor. In an effort to avoid the pitfalls of her predecessor, Zo will refuse to talk about any issued that could be considered sensible. Politics, religion, etc., are all excluded, the chatbot will try to change topics and ultimately refuse to engage. Furthermore, since her ability to understand context is not properly developed it will refuse to talk about these issues even if they are in a positive, or completely harmless context. As an example, if you tell Zo that a certain music was played at your bar mitzvah she will refuse to talk to you about it, because the word is filtered. If you vent that you are being bullied because of your religion she will give you the cold shoulder, while if you only tell her that you are being bullied (no reason provided) she will offer some support. Due to this fact it appears to have either deep hate or deep fear of the topics, which is, in itself, uncomfortable[311].

§ 1.6.5.3. The Racist AI: Part I – On Your Social Network

As we have mentioned before, datasets frequently suffer from a lack of diversity in their data. That may cause some regrettable mistakes for everyone involved, including the programmers who frequently have no malicious intentions.

Google Photos' auto-tagging feature had one such issue when it tagged the photo of two users as "gorillas" due to the colour of their skin. Google deeply apologised for the mistake. However, it seems that as with Microsoft and Tay, they were not able to address the root cause of the problem and to avoid similar situations just suppressed the gorilla and a few other tags. While not being able to tag gorillas automatically seems like a small

311 *See*, Tiago Sérgio Cabral, *"Robotics and AI in the European Union…*,135-146; "We really need to take accountability," Microsoft CEO on the 'Tay' chatbot", Charlie Moloney, accessed August 2, 2019, https://chatbotsmagazine.com/we-really-need-to-take-accountability-microsoft-ceo-on-the-tay-chatbot-2383ee83a6ba; "How Microsoft is Using AI to Tackle Fake News", James O Malley, accessed August 2, 2019, http://www.gizmodo.co.uk/2018/05/how-microsoft-is-using-ai-to-tackle-fake-news/; "Here Are the Microsoft Twitter Bot's Craziest Racist Rants", Sophie Kleeman, accessed August 4, 2019, https://gizmodo.com/here-are-the-microsoft-twitter-bot-s-craziest-racist-ra-1766820160;"Microsoft's neo-Nazi sexbot was a great lesson for makers of AI assistants", Rachel Metz accessed August 4, 2019, https://www.technologyreview.com/s/610634/microsofts-neo-nazi-sexbot-was-a-great-lesson-for-makers-of-ai-assistants/; "The Accountability of AI — Case Study: Microsoft's Tay Experiment", Yuxi Liu, accessed August 4, 2019, " https://chatbotslife.com/the-accountability-of-ai-case-study-microsofts-tay-experiment-ad577015181f; "Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter", Ellie Hunt, accessed August 5, 2019, https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter; "Microsoft's racist chatbot returns with drug-smoking Twitter meltdown"; Samuel Gibbs, accessed August 5, 2019, https://www.theguardian.com/technology/2016/mar/30/microsoft-racist-sexist-chatbot-twitter-drugs; "AI Robots Learning Racism, Sexism And Other Prejudices From Humans, Study Finds", Ian Johnston, accessed August 5, 2019, https://www.independent.co.uk/life-style/gadgets-and-tech/news/ai-robots-artificial-intelligence-racism-sexism-prejudice-bias-language-learn-from-humans-a7683161.html; "Microsoft's politically correct chatbot is even worse than its racist one", accessed August 5, 2019, https://qz.com/1340990/microsofts-politically-correct-chat-bot-is-even-worse-than-its-racist-one/.

price to pay for no racist mistakes, one has to wonder if it is not a weak solution for something that reveals a deeper problem (and a potential catalyst for a planet of the apes doomsday scenario). In fact, Flickr suffered similar problems in the past and so did Facebook (though in the latter case the issue was detected in testing) [312].

A study by Joy Buolamwini (founder of the Algorithmic Justice and computer scientist at MIT) and Timnit Gebru (a Microsoft researcher) showed that tools by Microsoft, FACE++and IBM are substantially less accurate at guessing the gender in women's photographs or photographs of people who are darker-skinned. The largest difference in accuracy is more than 20% for Microsoft and more than 30% for the remaining companies[313]. A subsequent study by Inioluwa Deborah Raji and Joy Buolamwini showed quite a leap in results for Microsoft, this time its maximum margin of error was 1,52% for dark-skinned females. FACE+ and IBM also improved, though they did not manage to achieve results similar to Microsoft. The study also included Rekognition, a type of facial recognition/analysis software that Amazon had been marketing to law enforcement. The results pointed out that it wrongly classified dark-skinned women 31,37% of the time. Amazon quickly reacted, considering that the researchers used the wrong tool offered by the software (analysis instead of recognition) and an outdated version, thus the conclusions had no scientific validity. Yet, a group of researchers from prestigious universities such as Harvard, Oxford and the IMT, and from industry leaders such as Google, Microsoft and DeepMind quickly got behind the study and called upon Amazon to stop selling Rekognition to law enforcement. The Amazon

---

[312] *See*, "Google says sorry for racist auto-tag in photo app", Jana Kasperkevic, accessed June 10, 2019, https://www.theguardian.com/technology/2015/jul/01/google-sorry-racist-auto-tag-photo-app; "Flickr faces complaints over 'offensive' auto-tagging for photos", Alex Hern, accessed June 10, 2019, https://www.theguardian.com/technology/2015/may/20/flickr-complaints-offensive-auto-tagging-photos; "Flickr Fixing 'Racist' Auto-Tagging Feature After Black Man Mislabeled 'Ape'", Michael Zhang, accessed June 10, 2019, https://petapixel.com/2015/05/20/flickr-fixing-racist-auto-tagging-feature-after-black-man-mislabeled-ape/; "Flickr's new auto-tags are racist and offensive", David Goldman, accessed June 10, https://money.cnn.com/2015/05/21/technology/flickr-racist-tags/; "Google apologises for Photos app's racist blunder", BBC News, accessed June 10, 2019, https://www.bbc.com/news/technology-33347866; "A major flaw in Google's algorithm allegedly tagged two black people's faces with the word 'gorillas'", Molly Mulshine, accessed June 10, 2019, https://www.businessinsider.com/google-tags-black-people-as-gorillas-2015-7; "When it Comes to Gorillas, Google Photos Remains Blind", Tom Simonite, accessed June 10, 2019, https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/; "Google 'fixed' its racist algorithm by removing gorillas from its image-labeling tech", James Vincent, accessed June 10, 2019, https://www.theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai; "Facebook is Finding Problems With Artificial Intelligence Too", Tom Simonite, accessed June 10, 2019, https://www.wired.com/story/facebook-finding-problems-artificial-intelligence-too/.

[313] *See*, Joy Buolamwini and Timnit Gebru, *"Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification"* Proceedings of Machine Learning Research (Conference on Fairness, Accountability and Transparency 81 (2018): 1-15; "Algorithmic Justice League", accessed June 10, 2019, https://www.ajlunited.org/.

board of directors did not agree, and shareholders voted down two proposals on the matter, one to prohibit the sale of the technology to governments and the second to study its impact on privacy and civil rights[314/315].

§ 1.6.5.4. The Racist AI: Part II – Imprisoning You

The Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is a tool developed by a Northpointe, Inc and used in certain US courts to predict recidivism by criminal offends[316]. When a defended is jailed, he/she is given a COMPAS questionnaire. The answers to said questionnaire are then fed to the algorithm which gives them a score of 1 to 10 in the categories of "Risk of Recidivism," "Risk of Violent Recidivism" and "Risk of Failure to Appear". A score of 1 to 4 would be considered low, 5 to 7 would be considered "medium" and 7 to 10 (high).

The ultimate decision always rests with the judges (or the jury), still a higher score on the "Risk of Failure to Appear" might influence the decision of whether a defendant is released pending trial or if he/she remains detained. A high "Risk of Recidivism" or "Risk

314 *See*, "Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products", Inioluwa Deborah Raji and Joy Buolamwini, accessed June 12, 2019, http://www.aies-conference.com/wp-content/uploads/2019/01/AIES-19_paper_223.pdf; "Amazon Is Pushing Facial Technology That a Study Says Could Be Biased", Natasha Singer, accessed June 12, 2019, https://www.nytimes.com/2019/01/24/technology/amazon-facial-technology-study.html; "Thoughts on Recent Research Paper and Associated Article on Amazon Rekognition", Matt Wood, accessed June 12, 2019, https://aws.amazon.com/blogs/machine-learning/thoughts-on-recent-research-paper-and-associated-article-on-amazon-rekognition/?; "Some Thoughts on Facial Recognition Legislation", Michael Punke, accessed June 12, https://aws.amazon.com/blogs/machine-learning/some-thoughts-on-facial-recognition-legislation/?; "Response: Racial and Gender bias in Amazon Rekognition — Commercial AI System for Analyzing Faces.", Joy Buolamwini, accessed June 12, 2019, https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced; "On Recent Research Auditing Commercial Facial Analysis Technology", Ali Alkhatib et al., accessed June 12, 2019, https://medium.com/@bu64dcjrytwitb8/on-recent-research-auditing-commercial-facial-analysis-technology-19148bda1832; "Microsoft improves facial recognition technology to perform well across all skin tones, genders", John Roach, accessed June 12, 2019, https://blogs.microsoft.com/ai/gender-skin-tone-facial-recognition-improvement/; "Mitigating Bias in AI Models", Ruchir Puri, accessed June 12, 2019, https://www.ibm.com/blogs/research/2018/02/mitigating-bias-ai-models/.

315 Even if they had succeeded, the votes were non-binding. Thus, the company could reject the suggestions and pursue business as usual. " Amazon shareholders reject facial recognition sale ban to governments", Zack Whittaker, accessed June 12, 2019, https://techcrunch.com/2019/05/22/amazon-reject-facial-recognition-proposals/; "Amazon shareholders support selling face recognition tech to police", France24, accessed June 12, 2019, https://www.france24.com/en/20190522-amazon-shareholders-support-selling-face-recognition-tech-police; "Amazon shareholders reject ban on selling face recognition software to police", Laurence Dodds, accessed June 12, 2019, https://www.telegraph.co.uk/technology/2019/05/22/amazon-shareholders-reject-ban-selling-face-recognition-software/.

316 Including Broward County, the State of Florida, the State of New York, the State of Wisconsin, and the State of California.

of Violent Recidivism" may result in someone not being granted conditional release. Besides, it could be argued that these two metrics could influence the length of the sentence itself and judges have, in fact cited COMPAS scores when sentencing.

It should be seen as extremely serious when an algorithm that may influence a person's freedom might be biased. When proper transparency and auditing tools do not exist for the tool, the issue intensifies.

A 2016 study by ProPublica drew attention to the fact that dark-skinned defendants were evaluated by the software as unfairly having a high degree of recidivism, while the evaluations of lighter-skinned defendants were reversed. In fact, dark-skinned defendants were labelled, wrongly, as likely to commit a crime in the future, nearly two times more frequently (45% against 23%). Similar statistics held for violent recidivism.

In general, the study considered that the algorithm was highly unreliable, only being able to predict 20% of violent recidivism and 61% of total recidivism.

Notwithstanding the discussion of whether this type of risk assessment is fair or compatible with fundamental rights, or even whether ProPublica's data is right[317], there is one thing we may be sure: if implemented without due attention it may perpetuate human bias contained within the data through the machine's decisions. If dark-skinned defends have been historically unfairly treated, it is only natural that an algorithm fed with the data of previous decisions will suffer from the same biases as those decisions[318/319/320].

---

317 Northpointe (obviously) disagrees but they are necessarily not alone in this position. *See*, "COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity", William Dieterich, Christina Mendoza anc Tim Brennan, accessed August 15, 2019, https://assets.documentcloud.org/documents/2998391/ProPublica-Commentary-Final-070616.pdf; "ProPublica's COMPAS Data Revisited", Matias Barenstein, accessed August 15, 2019, https://arxiv.org/pdf/1906.04711.pdf.

318 *See*, "Machine Bias", Julia Angwin et al., accessed August 15, 2019, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing; "Technical Response to Northpointe", Jeff Larson and Julia Angwin, accessed August 15, 2019, https://www.propublica.org/article/technical-response-to-northpointe.

319 Please note that the Council of Europe "adopted" the European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment, though it falls outside of the scope of this Thesis, and thus is not analysed, we find its principles vague, an absent binding power and a proper method of ensuring compliance its effectiveness would always be highly dubious. *See*, "European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment", Council of Europe, accessed October 1, 2019, https://rm.coe.int/ethical-charter-en-for-publication-4-december-2018/16808f699c

320 For more examples *see*, Nizan Geslevich Packin and Yafit Lev Aretz, "Learning Algorithms and Discrimination", *in Research Handbook on the Law of Artificial Intelligence*, Woodrow Barfield and Ugo Pagallo eds. (Cheltenham: Edgar Elgar Publishing, 2018), 88-113; Ronald Leenes and Silvia De Conca, "Artificial Intelligence and Privacy- AI Enters the House Through the Cloud", *in Research Handbook on the Law of Artificial*…, 280-306.

These stories share a common denominator. In all of them, it is not a certainty that the algorithms themselves were bad. Frequently algorithms are "poisoned" by the mistakes and biases that (sadly) already exist in our society. Insufficient data itself is also frequently a problem. But understanding what algorithms are "thinking" and why do they make certain decisions is key to avoid extending *ad eternum* our failures as a society through our machines. In addition, it is probably not enough for these decisions to be understood by programmers and AI specialists, the common user must have access to some information and know why he/she is being affected in this manner. Why is my CV not good enough to be select for this position? Why I was I considered at risk of violent recidivism?

Of course, having this knowledge is much more important when the decisions of an AI cause a significant impact on someone's life. Being insulted by an AI on Twitter should not be a pleasant experience, and may even have some psychological sequels but, at the end of the day, it seems far preferable to being imprisoned because the computer said so and did not like the colour of your skin. The problem is not restricted to AI that uses personal data in its functioning, but it is particularly dangerous there. There are other implications of bias, including in the fundamental rights and ethical realms, and even on the political and societal one.

Unfortunately, we cannot draw the reader's attention to all these stories and are sadly unable to explain all the consequences and aspects of algorithmic bias. However, we chose to put these stories before our arguments regarding the right to explanation so the importance of the thematic is understood by grounding on real life and current experiences.

§ 1.6.6. The Right to Explanation under the GDPR: Rights to Information and Access

Legal scholars have been debating whether the data subject has the right to full explanation or just to information and if said explanation or information should cover the specific decisions or just system functionality. In fact, the distinction between information and explanation is quite empty in this context, but we will address that in a moment.

The first two legal basis that we can find in the GDPR for the right to explanation are Articles 13 and 14. Under these dispositions the data subject has the right be informed of *"the existence of automated decision-making, including profiling, referred to in Article 22(1) and (4) and, at least in those cases, meaningful information about the logic involved, as well as the significance and*

the envisaged consequences of such processing for the data subject". There are a number of questions that must be addressed to understand this provision. First, not all automated decision-making is relevant, only when it falls within the scope of Article 22. That is to say *"a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her".* There is sound logic behind this choice, the legislator considers that there is a greater risk for the rights of the data subject in this type of processing. Clearly, if there are no relevant effects for the data subject's rights that greater risk disappears and, with it, the need for more protective rules. Imagine one of those "which character from X TV show are you?" Apps that collect the users' birth data, favourite colour and a few other miscellaneous (not sensible) information to provide them with an answer. There is automated decision-making, but provided that there is no further processing, its impact on the data subject is completely negligible and there is no reason to ask for the meaningful information about the logic involved or the significance and envisaged consequences of the processing[321].

In our preliminary questions, we already clarified what should be understood by based solely on automated processing and produces legal effects concerning him or her or similarly significantly affects him or her so we refer to the previous points on these matters.

§ 1.6.6.1. Under the Rights to Information and Access: Article 13 - Limitations

There is a reasonable degree of controversy surrounding the nature of the information (or explanation) that must be provided to the data subjects. Should it be information about the system functionality ("*logic, significance, envisaged consequences and general functionality*"' of the system) or the specific decision ("*the weighting of features, machine-defined case-specific decision rules, information about reference or profile groups*")? Should it be provided *ex ante* (before the automated decision-making) or *ex post* (after the automated decision-making)?

Should these divisions even be adopted? The abovementioned were constructed by Wachter and others. The Authors adhere to a very restrictive view of the right to

[321] The answer to the second and third question would be "there is no significance and no consequences are to be expected", the answer to the first one would be something in line with "if you answer c) more than Y times, you will be assigned character Z". In fact, while it could be slightly more complicated, the argument still stands, there is no reason for a more stringent rules. A different interpretation would be incoherent with the remaining provisions in the GDPR and with the legislator's intentions.

explanation – the original paper is called "Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation". Though we must point out that is not precisely what is defended by the Authors[322].

First, regarding the provision of meaningful information under the Right to Information we should separate Articles 13 and 14. For Article 13, information must be provided to the data subject when the personal data is collected. This moment is previous to the processing of the personal data for automated decision-making and thus, when data is collected the data controller still has no information about the specific decision. Thereby, it is only possible to provide information about the system functionality and not about a specific decision because there is none.

Andrew D. Selbst and Julia Powles rightly draw attention to the fact that, due to ML's deterministic nature, it should be possible to generate an explanation of the specific decision immediately when you have the input data. While technically, that is absolutely true, and you have to do no more than feeding the data to the model, reality is a bit more complicated. Even if the abovementioned Authors are right and their theory aligns with GDPR's logic there are two main issues with it. First, Article 13 establishes the following sequence of events: *a)* Collection and provision of information and then; *b)* further processing[323]. To apply Andrew D. Selbst and Julia Powles's method, provision of information would have to be moved to after further processing. Doing so would be in clear contradiction with the wording of Article 13 (and with recital 61 of the GDPR). Automated decisions also do not have to be made immediately when data is collected. For security, strategic or operational reasons, the system that collects data may not be the same as the one that makes the decisions and those may not be directly connected. Data may be pulled from one to the other to finish the operation. There is no guarantee that the same model is used for all data subjects (a different one may be used based on the region where the data subject lives for example). In fact, there may be a first model to decide which second model is adequate to deal with that specific data subject. Burdening the data controller with making an immediate decision would be extremely difficult and could, in fact, be counterproductive for operational or even security reasons.

With this in mind, meaningful information must still merit the qualification as "meaningful" even if it covers system functionality. The data controller does not have to

[322] Sandra Wachter, Brent Mittelstadt and Luciano Floridi, *"Why a Right to Explanation …,*79-99
[323] *See*, Andrew D. Selbst and Julia Powles, *"Meaningful information and the right to explanation"* International Data Privacy Law 7,4 (2017): 233-242.

disclose the full algorithm and, in fact, should avoid complex and non-understandable explanations about its inner workings. However, the data subject must be able to comprehend the basis for the decision that will be made about him or her. This could entail: *i)* explaining the data that will inform the algorithm in its decision *ii)* the most relevant aspects for arriving at a decision[324] , *iii)* standard scenarios to contextualize the data subject *iv)* measures to ensure that the algorithms used remain "fair, effective and unbiased" and; *v*) the right of the data subject to request human intervention and other relevant safeguards.

Regarding the significance and the envisaged consequences for the data subject, information about how the automated decision-making might affect the data subject using "meaningful and understandable, real, tangible examples of the type of possible effects should be given" in line with the WP29's Guidelines on Automated Decision-Making. The information should be provided for current and subsequent processing operations[325]. We must note that the real tangible examples are in line with our suggestion for standard scenarios and may be built jointly to meet both requirements, but are not the same. To meet the requirements of meaningful information about system functionality, you must build a standard scenario in which an individual with X characteristics provides Y data. With that data the model will produce Z result. For the significance and envisaged consequences what counts is the effect of the Z result on the individual. If the Z is a bad score, person A will not be provided with a loan.

§ 1.6.6.2. Under the Rights to Information and Access: Article 14 - Limitations

Article 14 shares most of the same limitations as Article 13, with one serious difference. Information according to Article 14 (when data is not obtained from the data subject) must be provided *"within a reasonable period after obtaining the personal data, but at the latest within one month, having regard to the specific circumstances in which the personal data are processed"*.

This means that, in theory, the automated decision could have already been reached before information is provided to the data subject. Still, we argue that such a scenario will

[324] Based on aggregated data of previous decisions if need be, or, if unavailable, the aggregated data from algorithmic testing.
[325] *See*, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826

only manifest exceptionally. There are two main reasons for this fact, the first being the legal basis under which personal data can be processed for automated decision-making and, the second, the concept of "reasonable period having regard to the specific circumstances".

Personal data can be processed for automated decision-making under the following legal basis: *a)* necessary for entering into or for the performance of a contract; *b)* authorised by EU or Member State law or; *c*) based on the data subject's explicit consent.

Explicit consent makes it impossible for data to be processed absent provision of information, since said consent would not be valid. Thus, the data controller may collect data but cannot process it for automated decision-making. If processing is compliant with the GDPR, there will never be a scenario where data is processed before information is provided grounded on the legal basis of consent,

Data that is necessary for entering into a contract or for the performance of a contract implies that there was a previous contact between the parties. Whether they already entered into a contract or are taking the necessary steps to do so, it is implied that both parties are aware of and willing to contract with each other. Thus, information should be provided in one of the previous instances where contact was established. Ideally, when the data controller first becomes aware that automated decision-making will be needed to enter or to perform said contract. Assuming that there is an available channel for communication between parties, processing of personal data for automated decision-making absence information also seems very unlikely or, at least, very rare.

Third, we have the option of the data processing being authorised by EU or Member State law, but even then, the same laws must also lay *"down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests"*. Under these suitable measures, there must be ones ensuring the provision of information to data subjects. In any case, it is expected that said measures will minimise the risk for data subjects' rights and interests, creating fewer issues than other types of processing. Still, this is the only of the three scenarios where we can imagine processing for automated decision-making happening without previous information being provided. If this happens the information to be provided after will need to be evaluated on a case-by-case basis, but it is possible that it may encompass more than information about system functionality.

There is also the concept of "reasonable period having regard to the specific circumstances". The one-month deadline in Article 14 is absolute and may be restricted if

it is not reasonable due to the specific circumstances of the processing[326]. How restricted depends on the case, but in its Guidelines on Transparency the WP29 points out that *"the principles of fairness and accountability under the GDPR require data controllers to always consider the reasonable expectations of data subjects, the effect that the processing may have on them and their ability to exercise their rights in relation to that processing, when deciding at what point to provide the Article 14 information"*. Data controllers must further be able to demonstrate the reasoning behind their decision. In fact, according to the WP29 *"wherever possible, data controllers should, in accordance with the principle of fairness, provide the information to data subjects well in advance of the stipulated time limits"*[327]. Thus, since this type of data processing activity is considered particularly "sensible" by the legislator, then, by nature, the processing will affect the data subject in a reasonably high manner thus making information yet more valuable, and taking into consideration other aspects that may arise in case, when possible, data controllers should inform data subjects before the processing activity (and automated decision-making) takes place.

Taking the abovementioned arguments into account, generally, information to be provided under Article 14 must follow the same rules as Article 13, that is, it must be on system functionality and have the aforementioned characteristics. When this proves false, and processing is done before the provision of information, a case-by-case analysis should be performed, but, in general, data controllers should then provide information about the specific decision also (in accordance with the rules about the right to access). If it is feasible and not disproportionate to do, the principle of transparency and fairness of processing, will generally require data controllers to disclose information before processing in this type of situations.

§ 1.6.6.3. Briefly: Valid Consent

For consent to be valid under the GDPR, it must be a *"freely given, specific, informed and unambiguous indication of the data subject's wishes"*. For consent to be freely given, there must be no clear imbalance of power. In its Guidelines on Consent the WP29 states that this will make public authorities unlikely to be able to rely on this legal basis for their data processing activities. Another situation where there might be a clear imbalance of powers

[326] Or when Articles 14 (3) point b) or c) are applicable.
[327] *See*, "Guidelines on transparency under Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=51025.

is an employment context. It is not the focus of the WP29's Guidelines, but it is important to point out that information asymmetry may equally give rise to a case of power imbalance. If the difference in information is excessive it may jeopardise the validly of consent as a whole. Thus, data controllers when using this legal basis should be extra careful when providing the data subject with the necessary information. Additionally, if information is not accessible or understandable to data subjects' consent may not be valid due to not being informed[328].

§ 1.6.6.4. Preliminary Question: Should Information Given Under Article 15 Be the Same as Under Articles 13 and 14?

As a preliminary question, we must point out that in its Guidelines on Automated Decision-Making[329] the WP29 states that the data controller should already have provided the data subject with the information [Article 15 information] in line with its Article 13 obligations. However, in our opinion this does not mean that the information provided under Article 13 and Article 15 must be exactly the same. In fact, information under Article 15 can, and should at times, be more detailed than under Article 13, as we will elaborate in a matter of moments[330].

As for the reasons serving as foundations to our conclusion, they are the following: in the Guidelines the WP29 is referring to compliance with the obligation of informing the data subject under Article 13 (and 14), namely the provision of information regarding:

a) the existence of automated decision making, including profiling;
b) meaningful information about the logic involved; and
c) he significance and envisaged consequences of such processing for the data subject.

As aforementioned, under the Right to Information this would entail the provision of information about system functionality and not about the specific decision. Still, reading

[328] *See*, "Guidelines on consent under Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=51030
[329] *See*, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826
[330] If information under Article 13 (and 14) is "A", information under the Right to Access will be "A + X".

the norm and WP29's opinion in such a manner is both artificial and unrealistic. The WP29 states the information should have been previously provided in line with Article 13 obligations but it does not say that said obligations are the same as Article 15 obligations. The Guidelines themselves address this fact saying that *"by exercising their Article 15 rights, the data subject can become aware of a decision made concerning him or her"*. Being made aware absent having the tools to understand would not provide the data subject with more than a (possible) illusion of control. Therefore, it is unlikely that when the WP29 refers to it, it does so in a superficial manner. Awareness should be understood as covering both awareness of the decision and awareness of the reasons for the decision.

If we take into account the principle of transparency, it is also clear that, since the data controller has new information that will be helpful for the data subject to exercise his or her rights, it should be disclosed.

This is not a change in the information to be provided under Articles 13 and 14. As established, information to be disclosed under these Articles is restricted by the timing in which it must legally be provided. Therefore, if the model (the system) does not change, neither does the information about it. However, the right to access is not restricted by these timing aspects and, therefore, more information can be provided[331].

Conceptually it would make no sense for the data subject to be provided with information about how the system works, the result of system's computations (the conclusions or inferences) and not about how the system arrived at those results in his/her specific case.

According to the WP29 *"the controller should provide the data subject with general information (notably, on factors taken into account for the decision-making process, and on their respective 'weight' on an aggregate level) which is also useful for him or her to challenge the decision"*[332]. Having this information is essential, but it is not enough to establish, for example, bias in algorithmic decisions and due to this fact, without specific information the data subjects will not be able to exercise their rights and the principle of fairness will not be realised (as we will explain below). Thereby, the Guidelines should not be read as restricting the provision of

[331] *See*, "Guidelines on transparency under Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=51025.
[332] "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826.

specific information, just establishing that general comparing information should also be provided.

§ 1.6.6.5. Under the Rights to Information and Access: Article 15

Rules will be slightly different under the right to access when compared to the ones regulating the right to information. In most cases, the processing for automated decision-making will have already taken place. As established above, information under the right to access does not have to be the same as information under the right to information and can be broader. Therefore, there is no technical reason to justify a restriction of explanation of the specific case to the right to access, in line with what we defended for the right to information. With this in mind, is there a right to explanation of the specific decision under the right to access or are any other reasons to argue that it does remain restricted to system functionality?

First, the division between information and explanation should be rejected as meaningless and artificial. As argued by Selbst and Powles *"if 'meaningful' is to have any substance, that appears on its face to be a move in the direction of explanation of some type*[333]. In fact, even information about system functionality to be useful for the data subject implies explanation (the data controller cannot just make available the technical guide for using the software/algorithm).

Providing information regarding the specific decision is, arguably, easier and less harmful for the interests of the data controller than providing general information about the system's inners workings. To provide system information, for example in a model that is based on deep learning, the data controller must provide information on a model which will evaluate hundreds if not thousands of parameters and that, in fact, may not be fully understood. To provide information on a specific decision the data controller "just" needs to provide information on the parameters (the weights) that were more relevant for that decision. If you are able to accomplish the first you will certainly be able to accomplish the second, but the opposite is not necessarily true. For data controllers the stricter interpretation offers a plethora of risks absent almost any advantage. Data controllers will have to build an explanation about the more challenging system functionality, and by not providing one relating to specific decisions, will be at the mercy of a possible (and quite

[333] *See*, Andrew D. Selbst and Julia Powles, *"Meaningful information*…, 233-242.

likely) restrictive interpretation by data protection supervisors and by the ECJ. Fact is that, absent a revolution in the manner in which the ECJ decides on data protection we would be surprised by an interpretation that required only information on system functionality. While this does not speak to the legal merits of the argument it is our belief that is an important disclaimer that data controllers should be aware.

Looking at the issue from another angle, is there a right to explanation of the specific decisions under the GDPR and, namely, under the right to access? To understand this issue, we must understand what the purpose of such a right is. The answer should be to ensure that the principles for the processing of personal data are complied with and that the data subject can exercise his or her rights under the GDPR (and under the EU's constitutional law).

Within these rights we may count access, erasure, rectification, restriction of processing, portability and right to object. However, the data subject also has the right to lodge a complaint with a supervisory authority (Article 78 GDPR), the right to an effective judicial remedy against a controller or processor (Article 79 GDPR) and the right to compensation or liability (Article 82 GDPR). Absent information about the specific decision it would be truly impossible for a data subject to prove damages and therefore pursue compensation in accordance with Articles 79 and 82 of the GDPR and Article 47 of the Charter. Even if a data processing operation is not fair, if its "unfairness" did not cause damage to a specific citizen it is unlikely that compensation can be granted. The absence of a specific explanation would hinder severely the capacity of data subjects to exercise these rights in cases of algorithmic discrimination. To accept such an interpretation would be entirely out of line with the general principles of the GDPR.

In addition, Recital 71 should not be ignored. Even if it has no binding authority and relates mainly to Article 22, it clearly demonstrates that the European legislator desired to enshrine a proper right to explanation of the specific situation. In fact, this interpretation avoids an unnecessary controversy on the nature of the adequate safeguards under Article 22. With a right to explanation under the previous Articles, the safeguard of explanation for a particular data subject is already ensured. It is our belief that this interpretation is the one that better reflects the legislator's thinking.

The information to be provided should be enough to allow the data subject to exercise his or her rights and should reflect the requirements of the WP29 for general information, but applied to the specific situation, namely the factors taken into account

for the decision-making process, and on their respective 'weight". Standard scenarios should still be provided to the data subject to allow the data subject to compare his/her situation with the generality of users.

As a final note, according to Recital 63, the right to explanation should *"not adversely affect the rights or freedoms of others, including trade secrets or intellectual property and in particular the copyright protecting the software. However, the result of those considerations should not be a refusal to provide all information to the data subject"*[334]. The situations where the right to data protection conflicts with rights of the data controller such as keeping trade secrets and intellectual property rights will have to be analysed case-by-case. This should not be seen, at all, like a (albeit non-binding) attempt to restrict data subjects' rights. In fact, the same conflict and result would arise if no recital existed. The wording of the recital actually appears to put the burden on the data controller, to find a manner to disclose the information without hurting its own rights. Of course, it also reinforces our considerations regarding how the controller does not need to disclose the source code or every single factor weighted by the algorithm in the decision (only the most important ones that are truly needed to guarantee the data subject's rights)[335].

§ 1.6.7. The Right to Explanation under the GDPR: Automated Individual Decision-Making

Having concluded that there is a right to an explanation of the specific decision under Article 15, we could have completely avoided the controversy around the specific safeguards of Article 22. However, there are good reasons not to do so. First, the wording of Article 22, in fact, reinforces our conclusions about Articles 13-15. Second, even if our previous arguments were not accepted the right to explanation of the specific decision is also contained within Article 22's safeguards and, as a precaution, we intend to cover all of the possible scenarios.

Article 22 states that when processing is based on explicit consent or into its necessity for entering or performing a contract *"the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to*

[334] The concept of legibility defended by Malgieri and Comandé can provide an adequate starting point for compliance by data controllers, since it meets and may even exceed the necessary requirements. *See*, Gianclaudio Malgieri and Giovanni Comande, *"Why a Right to*…, 243-265

[335] Doing so would actually create confusion in the data subject and hurt the principle of transparency itself.

obtain human intervention[336] *on the part of the controller, to express his or her point of view and to contest the decision"*. This provision is complemented by Recital 71 where it stated that *"processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision"*.

By accepting our interpretation of Articles 13-15, the suitable safeguard of obtaining an explanation in regards to the decision concerning him or her is, in fact, a legal requirement contained in both Articles 15 and 22. This position is in line with the interpretation contained in Recital 71, with the general principles of transparency and fairness in processing of personal data and with the rights of the data subject. It also explains why the legislator did not feel the need to expressly state in Article 22 something that is already a requirement under Article 15. It is probably the solution that best reflects the philosophy behind the GDPR, the line of interpretation followed by the ECJ, meanwhile managing to guarantee coherent application and not burdening the data controller with unrealistic requirements, such as asking for an explanation of the specific situation before a decision exists. Even if recitals are not binding, as explained in our section about their value in EU law, they are a highly valuable interpretative source and, the concept of meaningful information (and adequate safeguards also) are the type of open concepts where their guidance is most valuable. In our previous sections we also had the opportunity to rebuke argumentation that favoured ignoring the recital altogether

There is another key fact: in Article 22 it is stated point blank that data subjects have the right to express their point of view or to contest the decision. Explanation of the specific situation is clearly implicit in these ideas. Expressing their point of view should be interpreted as expressing their point of view in an informed manner, and it is only possible to express one's point of view in relation to a decision if one knows the arguments behind the decision. To think otherwise would be to concede that this right is no more than a placebo. The data subject can say that he/she does not agree but is not given "the weapons"[337] to mount a proper response. Contesting the decision should be interpreted broadly, as contesting the decision to the data controller, but also to a supervisory authority or court of law. This is in line with the provisions of the GDPR which give the data subject

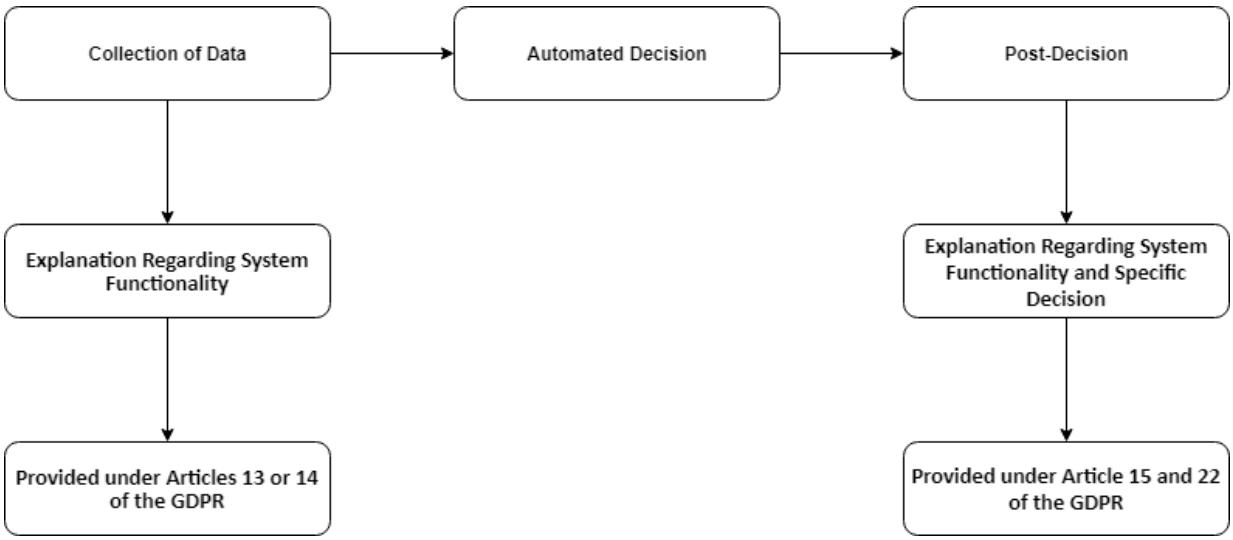
[336] We will study the question of human intervention below, for now we shall focus on the right to explanation.

[337] The WP29 is clear in saying the "*The data subject will only be able to challenge a decision or express their view if they fully understand how it has been made and on what basis".See*, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", WP29, accessed 15 June 2019,

all these rights. In fact, such diligences would not be possible (or at least could never succeed) without information. The aforementioned interpretation is clearly in line with our argumentation regarding the existence of a right to information regarding the specific situation contained within the right to access. Still, even if such a position is not adopted a right to information regarding the specific situation would still be contained within Article 22, without the need to apply or consider directly operative provisions of a recital. It needs no more than a functional and realistic interpretation that respects the general principles of the GDPR[338].

The diagram below presents a short summary of our position:

[338] The following works proved essential in writing sections 1.6.6. through 1.6.7..Thus, we opted to included them as general relevant sources and in the text of those sections to only use footnotes where there are direct citations. Sandra Wachter, Brent Mittelstadt and Luciano Floridi, *"Why a Right to Explanation …,*79-99; Andrew D. Selbst and Julia Powles, *"Meaningful information and the right to explanation"* International Data Privacy Law 7,4 (2017): 233-242; Gianclaudio Malgieri and Giovanni Comande, *"Why a Right to*…, 243-265; Lilian Edwards and Michael Veale, *"Slave to the Algorithm? Why a 'Right to an Explanation' is Probably not the Remedy You are Looking For"*, Duke Law & Technology Review 16,1 (2017): 18-84; Gianclaudio Malgieri, "Automated decision-making in the EU Member States: The right to explanation and other *"suitable safeguards"in the national legislation"*, Computer Law & Security Review (2019): in Press; " European Union regulations on algorithmic decision-making and a "right to explanation"", Bryce Goodman and Seth Flaxman, accessed August 17, 2019, https://arxiv.org/abs/1606.08813; Sandra Wachter, Brent Mittelstadt and Chris Russell, *"Counterfactual Explanations Without Opening the Black Box: Automated Decisions And The GDPR"* Harvard Journal of Law & Technology 31,2 (2018): 841-887; Mireille Hildebrandt, *"Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning"* 20,1 (2019): 83-122; "The Right to Explanation, Explained", Margot E. Kaminski, https://osf.io/preprints/lawarxiv/rgeus/download; Gianclaudio Malgieri, *"Trade Secrets v Personal Data: a possible solution for balancing rights"*, International Data Privacy Law 6,2 (2016): 102-116.
"Guidelines on transparency under Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=51025; "Guidelines on consent under Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=51030,"Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826; "Is there a right to explanation' for machine learning in the GDPR?", Andrew Burt, accessed June 17, 2019, https://iapp.org/news/a/is-there-a-right-to-explanation-for-machine-learning-in-the-gdpr/.



The diagram is a generalization and there are exceptions. One exception can arise when information under Article 14 is provided after the decision is made. However, as explained above that will be the exception and not the rule. Furthermore, both Articles 15 and 22 by themselves provide a right to explanation of the specific decision, a stronger argument can be made by joining the provisions of both.

The solution for the issue of black box algorithms and lack of human explanation may be technological one. Development of explainable XAI solutions where the algorithm itself is able to produce and explanation for its decisions and that explanation is understandable by humans has been gaining traction and can, indeed, solve the conundrum. Of course, this does depend on how advanced the technology for providing explanation is and if the explanation meets the current requirements abovementioned and future requirements that will certainly appear on AI-specific legislation. While XAI appears to have the potential to be a cornerstone of the European Trustworthy AI strategy and private companies such as IBM are already investing in it, if the EU really is serious about taking a leading role as the standard-setter for AI, it must provide adequate funding for XAI development and the necessary research, testing and development conditions for this

technology to reach a degree of advancement where it can both provide the adequate explanations and results similar in quality to "black box AI"[339/340].

§ 1.7. Automated Decision-Making and the Right to Obtain Human Intervention

The scope of the right to obtain human intervention under Article 22, must also be studied. This is a mandatory safeguard that must be implemented by the data controller, but that, if interpreted in the wrong way, may be impossible to implement or, at least, have a serious chilling effect on innovation.

More and more machines will be able to make decisions that will fall within the scope of Article 22. It is not adequate to expect that human reviewers will be able to take into account into their review process the same quantity of data processed by algorithms. A practical example, the Commission Implementing Regulation 2019/947/EU of 24 May 2019 on the rules and procedures for the operation of unmanned aircraft (Drones Regulation) establishes that to fly certain categories of drones, remote pilots may have to take an online training course and to have passed the online theoretical knowledge examination. Said examinations will probably be under Article 22 (2), point b) but the argument still stands[341]. These will probably be multiple-choice and automatically corrected by computer[342]. Obviously, we are not talking about highly advanced ML-based software agents, but it is still automated decision-making. If a significant number of examinees decide to exercise their right to human intervention should the data controller have to

---

339 Unfortunately, the ML-techniques that are currently achieving the most impressive results such as Deep Learning (see our section on Neural Networks) are also the most difficult to explain and, therefore, to develop an XAI solution for. Thus, adequate funding and conditions for research seem to be necessary to overcome this obstacle.

340 *See*, "Introducing AI Explainability 360", Aleksandra Mojsilovic, accessed August 20, 2019, https://www.ibm.com/blogs/research/2019/08/ai-explainability-360/; Liya Ding , *"Human Knowledge in Constructing AI Systems – Neural Logic Networks Approach towards an Explainable AI"* Procedia Computer Science 126 (2018) 1561–1570; "Designing Theory-Driven User-Centric Explainable AI", Danding Wang et al., accessed August 20, 2019, https://www.ashrafabdul.com/pdf/xai-framework-preprint-chi2019.pdf; "Explainable AI Driving business value through greater understanding", Chris Oxborough et al., accessed August 20, 2019, https://www.pwc.co.uk/audit-assurance/assets/explainable-ai.pdf; Robert R. Hoffman, Gary Klein and Shane T. Mueller, *"Explaining Explanation For "Explainable AI"* Proceedings of the Human Factors and Ergonomics Society 2018 Annual Meeting 62,1 (2018): 197-201.

341 In any case, Recital 71 seems to support the idea that human intervention also needs to be one of the appropriate safeguards under Article 22(2), point b).

342 An example based on the performance of a contract could be an e-learning platform who uses multiple-choice tests to evaluate its students. If they fail, they will not be able to receive the certificate as provided in the contract. Another option would be a contest, where to pass to the second-stage contestants need to fill a multiple-choice test, the final reward is a large sum of money. The legal basis of this last case would be consent.

provide someone who manually reviews each and every question in each theoretical exam? The answer has to be negative.

The WP29 considers that *"any review must be carried out by someone who has the appropriate authority and capability to change the decision. The reviewer should undertake a thorough assessment of all the relevant data, including any additional information provided by the data subject"*[343]. Any relevant data should be read as elements brought to light by the data subject, that could be relevant to the decision. In our example, after providing information to the data subjects regarding both system functioning and specific decisions, the data subject may point out that one of the questions is not being corrected properly, or that there is a second valid interpretation. The software probably will not understand this nuance, but a human will. Still, if the human suspects, given the information that he/she received from data subjects that the algorithm is biased or is not processing personal data in accordance with the rules of the GDPR, the reviewer should have the ability to request an audit. Obviously and, as stated above, the person reviewing should be able to detect errors, at least, at the level of someone who is reasonably familiarised with the program and is reasonably diligent and, of course, have the power to override the decision.

For a slightly more advanced and ML-related explanation, let us imagine that a video hosting website is using an algorithm to ensure that the videos uploaded to the website comply with its terms of service, for example videos have to be family-friendly and contain no nudity or curse words. This may have grave consequences for the data subject, since losing the ability to upload videos may also mean not being able to monetise them or having his/her account banned permanently. According to our previous argumentation, when one of the videos is rejected the data subject may ask, not only for an explanation about the system functioning but also about the specific decision. With this information the data subject will know why the video was rejected and will be able to contest and ask for human intervention. The software detected a forbidden word in minute 1:56 and that is why it was banned. By looking at it the data subject may understand that, in fact, the curse word to which the software refers is used to pass an anti-bullying message and give that information to the human reviewer. The reviewer aware of the nuances can now approve the video and override the machine's decisions.

[343] *See*, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826

This is another reason why our interpretation regarding the right to explanation for automated decision-making is important. Even if it, at first glance, appears to favour data subjects against data controllers it is not as black and white as that. In fact, it shares responsibility in a balanced manner. If we only provided explanation about system functionality, the data subject would have no way of expressing his/her opinion or contesting in an informed manner. Further, when doing so and asking for human intervention the data subject would only be able to express general disagreement. The consequence is that the entire burden of the review process would fall on the part that has all the information: the data controller. If we give information to the data subject, not only are we able to ensure a more effective and successful exercise of rights, we are also able to lessen the burden on data controllers who will have to review only "relevant" data and keep an eye out for possible defects and bugs in the algorithm.

§ 2. Shortly: The Charter of Fundamental Rights of the European Union

One quick note on the applicability of the Charter to the development and deployment of AI. If EU-wide legislation is implemented regulating AI or when applying current legislation such as the GDPR the scope of application of the Charter of Fundamental Rights of the European Union will be triggered, according to Article 51 of this instrument. In this scenario, at least, Member States would be bound to protect the rights enshrined in the Charter in their implementation and application of the law. The consequences may be wider if we accept horizontal effect to these fundamental rights, either directly or through considering them as general principles of EU Law[344].

Whatever the case, those working on AI-development would do good to keep in mind the principles in the Charter and labour to build AI which complies with these principles as, in one form or another, there is always the possibility of them being called upon. Furthermore, some of these principles will be highly influential and reflected into the principles enshrined in any AI-specific legislation from European source or may even be integrated directly through an article that States that AI to be developed or deployed in the European Union must comply with the principles and rights contained within the Charter and remaining European Constitutional Law.

[344] Alessandra Silveira, *Princípios de Direito da* …, 103ff.

Chapter II – AI-enabled Devices and Services and Consumer's Rights: A Short Introduction to the EU's new Legal Framework

In this chapter we will study the changes introduced to the EU's legal framework by the recently approved Directive 2019/771/EU of the European Parliament and of the Council of 20 May 2019 on certain aspects concerning contracts for the sale of goods, amending Regulation 2017/2394/EU and Directive 2009/22/EC, and repealing Directive 1999/44/EC (hereinafter, Sale of Goods Directive) and Directive 2019/770/EU of the European Parliament and of the Council of 20 May 2019 on certain aspects concerning contracts for the supply of digital content and digital services (hereinafter, "Digital Services Directive").

It is not our intention to pursue a thorough assessment of the impact of AI across the entire world of EU's consumer protection legal framework. Of course, space being limited we opted to restrict our analysis to legislation that will have the most impact on the development and adoption of AI-enabled technologies. Being that the Sale of Goods Directive and Digital Services Directive were written with new generation technologies in mind and, clearly, specifically AI, it is our opinion that the success of their provisions will be extremely important in setting up a relationship of trust between consumers and technology and in ensuring that sellers provide durable and trustworthy technology to European consumers.

§ 1. The Sale of Goods Directive

Both the Digital Services Directive and the Sale of Goods Directive are part of a package intended to modernise and adapt consumer protection to our current market. Directive 1999/44/EC of the European Parliament and of the Council of 25 May 1999 on certain aspects of the sale of consumer goods and associated guarantees (hereinafter, "Directive 1999/44/EC") is highly minimalistic, currently not adapted to our technological world and its minimum harmonisation approach proved not to be adequate for the single market. As a result, rules on consumer protection across the European Union significantly differ on matters as important as the hierarchy of remedies (or lack of thereof) and or the period in which the lack of conformity is presumed to have existed at the time of delivery that varies widely between Member States (it may go as low as 6 months).

With rules that vary to such a degree, traders are not able to fully enjoy the benefits of the single market because they do not know what rules will expect them when selling to a different State (and they may be very different), neither consumers feel the peace of mind necessary to purchase products in another Member State because they are not sure that they will enjoy a high level of consumer protection. The issue is especially problematic in the Digital Single Market. The new Sale of Goods Directive (a maximum harmonisation Directive) tries to tackle this challenge with higher default protection, fewer exceptions where Member States may opt for a different set of rules (and when they can the scope is more restricted) and a higher degree of adaptation to the Digital Single Market.

The Sale of Goods Directive directly references its applicability to the sale of goods including digital elements. According to Article 2 (and Recital 14) of the Sale of Goods Directive, goods including digital elements are *"any tangible movable items that incorporate or are inter-connected with digital content or a digital service in such a way that the absence of that digital content or digital service would prevent the goods from performing their functions"*. This is the case with "smart" products and appliances such as smartphones, smart TVs and smart fridges etc. These types of products may be AI-enabled and, in fact, may even be self-learning. Imagine a smart fridge that "learns" your favourite foods and stores, suggests promotions in accordance and even allows you to do your shopping through it (or does it for you!).

Robots will also fall straight into this Directive's scope of application. This will be true for our current Robots such as house cleaning robots and to future and more advanced technologies in the same line.

Interestingly the seller is still liable for *"o the consumer for any lack of conformity which exists at the time when the goods were delivered and which becomes apparent within two years of that time"* but, in the case of *"the case of goods with digital elements, where the sales contract provides for a continuous supply of the digital content or digital service over a period of time, the seller shall also be liable for any lack of conformity of the digital content or digital service that occurs or becomes apparent within two years of the time when the goods with digital elements were delivered. Where the contract provides for a continuous supply for more than two years, the seller shall be liable for any lack of conformity of the digital content or digital service that occurs or becomes apparent within the period of time during which the digital content or digital service is to be supplied under the sales contract"*(Article 10)[345]. Any lack of conformity that appears within a year of the time when the goods were delivered is presumed to have existed at the time. Member States may replace this period with a longer

[345] Member States may introduce longer time limits.

one, with two years being the limit. For goods with digital elements where there is a continuous supply of the digital content or service, the rules of the burden of proof follow the rules for liability of the seller.

The legal remedies repair, replacement, reduction of price and termination are now clearly applicable due to issues with the essential incorporated or inter-connected digital content or digital service. Furthermore, there is now a clear obligation by the seller to provide updates, including security updates. In most cases, updates should be provided, at least, until the period in which the seller is liable for any lack of conformity ends (Recital 31 and Article 7). Updates are different from upgrades (as in new versions of the system). Still, it is interesting to see that the fact that updates are not provided may give rise to a situation of lack of conformity and use of the legal remedies by the consumer.

The Sale of Goods Directive must be transposed by Member States 1 July 2021 and its provisions have to applied from 1 January 2022[346/347].

§ 2. The Digital Services Directive

The Digital Services Directive complements the Sale of Goods Directive by regulating Digital Content and Services when not part of a good with digital content. The scope of application of the Digital Services Directive includes, but is not limited to, *"computer programmes, applications, video files, audio files, music files, digital games, e-books or other e-publications, and also digital services which allow the creation of, processing of, accessing or storage of data in digital form, including software-as-a-service, such as video and audio sharing and other file hosting, word processing or games offered in the cloud computing environment and social media"* (Recital 19). It is plain to see plenty of AI-enabled technologies and consumer-grade software agents will fall within the scope of this Directive.

[346] Its provisions shall not apply to contracts concluded before 1 January 2022.

[347] "New perspectives on sale of consumer goods – maximum harmonization and high protection of consumers as a condition for the further development of cross-border trade in single market", Maria João Pestana de Vasconcelos, accessed May 30, 2019, https://officialblogofunio.com/2019/05/13/new-perspectives-on-sale-of-consumer-goods-maximum-harmonization-and-high-protection-of-consumers-as-a-condition-for-the-further-development-of-cross-border-trade-in-single-market/; "All for one and one for all - EU consumer remedies unite", Edward Turtle and Evangelia Nitti, accessed May 30, 2019, https://www.lexology.com/library/detail.aspx?g=15152d7d-462d-4e07-b8e7-fde75d677c2e; Jorge Morais de Carvalho, *"Sale of Goods and Supply of Digital Content and Digital Services – Overview of Directives 2019/770 and 2019/771"* Journal of European Consumer and Market Law 8,5 (2019): 194-201.

A significant exception is that open-source software and free software will be exempt from the Directives requirements. This is in line with the EU's intention of promoting open-source software development. In this context, and in a clear attempt to rein in the business model of certain large internet giants, free means free *stricto sensu*. That is to say that to fall outside of the scope of the Directive, data processing in free software must be restricted to what is needed for the purpose of improving the security, compatibility or interoperability of that specific software[348].

In this regard, while conserving that "personal data is not a commodity" the Digital Services Directive introduces provisions to protect consumers in contracts where services are provided in exchange for personal data. In fact, the GDPR restricts this practice but it does not seem to establish that it is not permissible at all, if the adequate measures are implemented. Still, data controllers should be careful when deciding the legal basis for data processing operations. Even if there appears to be some ground in the Directive for data as counter-performance businesses, data protection considerations will have to be applied.

Most of the times, data controllers will not be able to argue that data is necessary for the performance of a contract. In accordance with the WP29's legitimate interests Opinion, the necessity for the performance of a contract should be *"interpreted strictly and does not cover situations where the processing is not genuinely necessary for the performance of a contract, but rather unilaterally imposed on the data subject by the controller. Also the fact that some data processing is covered by a contract does not automatically mean that the processing is necessary for its performance"*[349]. The EDPB's "Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects" share the same tone. The Guidelines further add that *"Contracts for digital services may incorporate express terms that impose additional conditions about advertising, payments or cookies, amongst other things. A contract cannot artificially expand the categories of personal data or types of processing operation that the controller*

[348] Or in open source software for the matter. Though the issue in open source software would always be minor in comparison with free but closed source software. If an open source software was collecting data for unrelated purposes, the most likely situation is that a *fork* removing the data collection would quickly appear giving users a choice.

[349] *See, Opinion* 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC", WP29, accessed June 8, 2019, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf.

needs to carry out for the performance of the contract"[350/351]. Of course, performance of a contract may still be the correct and proper legal basis for data processing. In fact, AI-enabled devices that "learn" to personalise themselves in accordance with the consumer's behaviour and where that characteristic is key for their functioning and for the consumer's interest in the service/content will, probably, be able to use this legal basis, whether they are free or not.

However, in the remaining cases and for purposes that are not essential to the provision of the service or to make the content available, the data controller should rely either on consent or in the legitimate interests pursued by the data controller or a third party. As an example, if you are providing a video streaming service where recommendations are a key aspect you will probably be able to use performance of a contract as a legal basis. However, if you intend to use the data for targeted advertising or share it with commercial partners you will have to find a different and adequate legal basis. Additionally, not all data can be collected, you will have to have a justification regarding why do you need certain data for your recommendation system. You may not need to know someone's nationality for your recommendation system, but their address and preferred language may prove to be more relevant[352].

---

350 The European Data Protection Supervisor (EDPS) was also called to give its opinion on the Directive when it still a proposal. The Regulator was highly critical of its provisions considering that *"there might well be a market for personal data, just like there is, tragically, a market for live human organs, but that does not mean that we can or should give that market the blessing of legislation"* and establishing the differences between paying with money and "paying with data" where the consumer *"[whereas] the consumer is aware of what he is giving when he pays with money, the same cannot be said about data. Standard contractual terms and privacy policies do not make it easy for the consumer to understand what is precisely made with the data collected about him/her". See*, "Opinion 4/2017 on the Proposal for a Directive on certain aspects concerning contracts for the supply of digital content", EDPS, accessed June 8, 2019, https://edps.europa.eu/sites/edp/files/publication/17-03-14_opinion_digital_content_en.pdf.

351 *See*, "Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects", EDPB, accessed October 27, 2019, https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_en.pdf

352 For what is worth, this is just a theoretical example. In fact, if you are charging a value for your service you will probably need the address anyway, for tax purposes (but unless it is also needed for the functioning of your service you then will only be able to use it in that manner). In our example, we considered that preferred language is key to give the person content that is spoken, dubbed or subtitled in that language and that the address may help you provide content that is generally more popular in a certain region and avoid, at least, until you have a higher degree of information that allows you to give a recommendation properly targeted to the individual, hurting the cultural sensibility of the person. The same argument could be made about nationality, but we find it weaker in this context and if you weighted the situations where you would be collecting data that is not useful against the ones where it is useful, the damage to the privacy of many would be more relevant than the positive results for a few, and the principle of minimization would be applicable. Nevertheless, if you had a statistical model or some other information that made a strong argument for collecting nationality or other type of personal data if might be defensible. It is a really casuistic issue.

Still, both legal basis have limitations, consent must be valid in accordance with the GDPR, and that also means that it must be freely given. Therefore, if the data subject cannot deny consent absent negative consequences its validity is doubtful. A possible solution is to give a reasonably priced alternative and a free alternative depending on the provision of some data to the data subject, and let him/her choose[353]. The use of legitime interests depends on a balance test that must be conducted in an impartial manner and is forbidden where, for example, the specific processing triggers the application of either Article 9 or Article 22 of the GDPR[354].

For single acts of supply of digital contents and services, the trader shall be liable for any lack of conformity for a period of non-less than two years. The burden of proof shall rest on the trader for one year[355]. When there is a continuous supply over a period of time, the trader will be liable for a lack of conformity *"that occurs or becomes apparent within the period of time during which the digital content or digital service is to be supplied under the contract"*. The trader shall also be burdened with proving that lack of conformity did not exist for the abovementioned period of time.

Updates have to be provided in accordance with the contract and during the time that is needed to keep the digital content and service in conformity. The timeframe within which updates have to be provided mirrors the rules of the Sale of Goods Directive.

Remedies for lack of conformity are in line with the ones enshrined in the Sale of Goods Directive and failure to supply may, at the limit, lead to termination of the contract. *"The consumer shall be entitled to terminate the contract if the modification negatively impacts the consumer's access to or use of the digital content or digital service, unless such negative impact is only minor"* (Article 19).

---

353 See the decision by the Austrian Data Protection Supervising Authority where it considers that a mechanism where the data subject either consents to analytic or advertising cookies being installed in its machine or has to pay a value for the subscription of a newspaper is lawful. The issue at hand is related, firstly, to the e-Privacy Directive, but since said Directive applies the concept of consent under the GDPR (which was analysed in case) the conclusions are valid for GDPR-only cases. *See*, "Austria: "Cookie Walls / Paywalls" Hybrids Are Permissible?", Privacy Matters: DLA Piper's Global Privacy & Data Protection Resource, accessed July 5, 2019, https://blogs.dlapiper.com/privacymatters/austria-cookie-walls-paywalls-hybrids-are-permissible/.

354 You may find a deeper (and highly interesting) analysis on the interplay between the GDPR and the Digital Services Directive see Maria de Almeida Alves, *"Directive on certain aspects concerning contracts for the supply of digital content and digital services & the EU data protection legal framework: worlds colliding?"*, UNIO EU Law Journal 5,2 (2019): in press.

355 Unlike with the Sale of Goods Directive, this timeframe does not appear to extensible by Member States.

Importantly, the Directive does not restrict Member States ability to regulate liability claims against persons different from the trader, including the developer (Recital 13).

The Digital Services Directive must be transposed by Member States 1 July 2021 and its provisions have to applied from 1 January 2022[356/357].

Chapter III – Product Liability

§ 1. The Product Liability Directive

§ 1.1. Current Framework

Unlike in the fields of data protection (and arguably even consumer protection stricto sensu) the legal regime regulating producer's liability for defective products does not arise from a recent legislative effort. In fact, Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products (hereinafter, "Product Liability Directive")[358] is a legacy from a time where integration in the EU was not as deep as it is today, Council supremacy still reigned supreme and the Digital Single Market was still a remote idea, or not even that. The Product Liability Directive's intention is not to fully harmonise the European legal framework and space for derogations through national law

[356] The provisions of the Digital Services Directive shall apply to the supply of digital content or digital services which occurs from 1 January 2022 with the exception of Articles 19 (modification of the digital content or digital service) and 20 (right to redress) which shall only apply to contracts concluded from that date (Article 24).

[357] "Latest EU directives strengthen the protection of consumers in the digital world", Wolf Theiss, accessed June 5, 2019, https://www.lexology.com/library/detail.aspx?g=689432a8-4c62-4823-a011-96d8469dd662; "New rules on contracts for the supply of digital content and digital services", Cătălin Grigorescu and Diana Gavril, accessed June 5, 2019, https://www.bpv-grigorescu.com/publications/legal-tax-alerts/new-rules-on-contracts-for-the-supply-of-digital-content-and-digital-services/; "Not content with digital – new protections for EU consumers of digital content and services", Edward Turtle and Evangelia Nitti, accessed June 5, 2019, https://products.cooley.com/2019/03/15/not-content-with-digital-new-protections-for-eu-consumers-of-digital-content-and-services/; "Regulating the Economic Impact of Data as Counter-Performance: From the Illegality Doctrine to the Unfair Contract Terms Directive", Philipp Hacker, accessed June 13, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3391772&download=yes; *"Consumers, Liability, and the Online World"* , J.E.J Prins, accessed September 1, 2019, https://pure.uvt.nl/ws/portalfiles/portal/548857/CICT_12_2_05LORES.pdf.

[358] As amended by Directive 1999/34/EC of the European Parliament and of the Council of 10 May 1999 amending Council Directive 85/374/EEC on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products.

is quite broad in it, including, we argue excessively broad in provisions that seem core to the legal instrument.

In accordance with the Product Liability Directive the general rule in EU Law is strict liability of the Producer for damages caused by in the defects in the product[359]. National rules providing for contractual or non-contractual liability arising from damages of the same nature remain unaffected by the Directive, meaning that it is possible for the injured party to choose between the remedies offered by the different regimes or even request compensation under a combination of them, as he/she deems fit, within the limits of the law.

Producer, in accordance to the Directive is *"means the manufacturer of a finished product, the producer of any raw material or the manufacturer of a component part and any person who, by putting his name, trade mark or other distinguishing feature on the product presents himself as its producer"*. To correctly understand this definition, we should divide it in four parts. A Producer can be: *i)* the manufacturer of a finished product; *ii)* the manufacturer of any component part; *iii)* the producer of any raw material and; *iv)* any person, who presents himself as the product's producer. This aim of this provision is to act as an all-catching net, ensuring that anyone within the production chain can be held liable unless an exception applies (below) and giving the injured party plenty of choice of against whom to seek damages. For AI-enabled devices and AI-enabled robots, a broad notion of Producer can be particularly relevant. It is infrequent for an electronics company to produce every component part of the final product. It is already possible to see this in the design of modern computer and smart phones (which are, arguably, already AI-enabled devices). The chip on a modern computer will be produced by one company, another company produces the GPU, and the hard drives, and the screen etc. Some companies[360], may be able to produce a product from the ground up, but those are the exceptions[361]. With the development of more advanced and

[359] The injured party must prove the damage, the defect and the causal relationship between defect and damage.

[360] Samsung being a notable example. Even in Samsung's case, when the company needs and operative system it will generally go to a third-party: Android for mobile OS, and Windows for PC. Still, the company does have the capability and, in the past, invested in its own mobile operative system, albeit with limited success. For the PC, solutions are not hard to come by, through developing a Linux alternative (or using one those already available) for example. However, in practice, the consumer's preference (for various reasons) appears set on the current abovementioned combinations making any alternatives not economically viable.

[361] Apple invested heavily in research and development to start producing its own chips for the iPhone, and managed to do so with success. Nevertheless, it still needs to buy plenty of external components, as screens and memories (notably from its rival Samsung). Mac Computer's CPU's are generally bough from Intel, GPU's from Intel or Nvidia (higher-end computers get Nvidia chips while, lower-end ones are equipped with Intel integrated chips), hard drives and ram and rom memory from companies such as Samsung etc.

specialised AI-enabled products, this trend is unlikely to cease, and different entities will produce different parts of the final product. In addition, it is doubtful that any (technological) company by itself will ever produce both components and raw materials (or raw materials at all). In addition, the Product Liability Directive adds an extra layer of protection where products are imported to the EU, setting forth that the importer of a product *"for sale, hire, leasing or any form of distribution in the course of his business shall be deemed to be a producer"* and bear the same responsibilities as a Producer. Finally, when no Producer can be identified, every supplier in the supply chain can be called upon to answer as a Producer, unless the supplier is able to inform the injured party *"within a reasonable time, of the identity of the producer or of the person who supplied him with the product"*[362].

When more than one person is liable for damages under the rules of the Directive (which may happen with some frequency in our subject-matter), they will be liable both jointly and severally, without prejudice to national legislation that may provide the entities with solutions concerning the rights of contribution and recourse.

The concept of defective products is fleshed out in Article 6 of the Product Liability Directive and takes into consideration the level of safety which a person is entitled to expect. This expectation may vary in accordance with details that include: *a)* the presentation of the product, including expectations created by advertising; *b) "the use to which it could reasonably be expected that the product would be put"* and; *c)* the time when the product was put into circulation, an evolving concept, meaning a manufacturer can be considered liable for a defect for which it would not be liable in years past.

The concept of damages, by default, does not include non-material damages. Member States may put forth provisions establishing the contrary, though. Damages for the destruction of property have a lower threshold of 500€ to be compensable[363]. The Producer is only liable for damages to property when the item of property: *i) "is of a type ordinarily intended for private use or consumption"*, and *"was used by the injured person mainly for his own private use or consumption"*.

The rights of the injured person are extinguished three years after him/her "became aware, or should reasonably have become aware, of the damage, the defect and the identity of the producer". Notwithstanding, these rights shall equally be extinguished *"10 years from the date on which the producer put into circulation the actual product which caused the*

[362] This rule is also for imported products if they do not indicate the identity of the importer to the Union, even if the name of the Producer is identified.
[363] For reference, one ECU is equal to one Euro.

damage, unless the injured person has in the meantime instituted proceedings against the producer". Member States may also limit the amount for a which a Producer will be liable to 70€ million.

The Producer can avoid liability by proving that one of the requirements of Article 7 of the Product Liability Directive is fulfilled. One requirement that may be particularly relevant and AI development and deployment is Article 7, point e), stating that the Producer will not be held liable if *"the state of scientific and technical knowledge at the time when he put the product into circulation was not such as to enable the existence of the defect to be discovered"*[364] (hereinafter, "The State-of-the-Art Defence". The State-of-the-Art Defence is limited though, and its introduction in the Product Liability Directive bears every sign of having been the result of a difficult compromise. Firstly, Member States may, at will, completely remove the State of the Art Defence from their national frameworks and second the EC was mandated to submit a report regarding the effect that of the State-of-the-Art Defence and the exceptions to it created by Member States had on the single market, with the option of removing it altogether. The State-of-the-Art Defence survived this procedure. Nevertheless, as we will see below, there are more than a few issues surrounding it and its application that may get especially complex for AI-enabled devices[365].

---

[364] Additionally, the Producer will not be held liable when: *a) he did not put the product into circulation*; *b) that, having regard to the circumstances, it is probable that the defect which caused the damage did not exist at the time when the product was put into circulation by him or that this defect came into being afterwards; c) that the product was neither manufactured by him for sale or any form of distribution for economic purpose nor manufactured or distributed by him in the course of his business; d) that the defect is due to compliance of the product with mandatory regulations issued by the public authorities or; e) in the case of a manufacturer of a component, that the defect is attributable to the design of the product in which the component has been fitted or to the instructions given by the manufacturer of the product."*

[365] *See*, "Product Liability Through the Prism of EU Law", Kevin Rihtar, accessed September 1, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2391518; Judgment of the ECJ of 10 May 2001, *Henning Veedfald v Århus Amtskommune*, Case C-203/99, ECLI:EU:C:2001:258; "(Third) Report From The Commission To The European Parliament, The Council And The European Economic And Social Committee on the Application of the Council Directive on the approximation of the laws, regulations, and administrative provisions of the Member States concerning liability for defective products (85/374/EEC)", European Commission, accessed 15 August, 2019, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0246&from=en; "Report From The Commission To The European Parliament, The Council And The European Economic And Social Committee: Fourth report on the application of Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products amended by Directive 1999/34/EC of the European Parliament and of the Council of 10 May 1999", accessed 15 August, 2019, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52011DC0547&from=EN

§ 1.2. Potential Shortcomings on AI Regulation

§ 1.2.1. Is AI a Product or Service?

An interesting comparison can be made between the Digital Services Directive and Sale of Goods Directive which will regulate liability for traders/sellers[366] regarding goods and services that may be AI-equipped and enabled. The Sale of Goods Directive is applicable to goods meaning *"any tangible movable items; water, gas and electricity where they are put up for sale in a limited volume or a set quantity"*, including goods with digital elements, which are defined as "*any tangible movable items that incorporate or are inter-connected with digital content or a digital service in such a way that the absence of that digital content or digital service would prevent the goods from performing their functions*". For digital content and digital services that do not fall into the scope of the Sale of Goods Directive the legislator created the Digital Services Directive. Digital content shall be understood as *"data which are produced and supplied in digital form"* and digital services a *"service that allows the consumer to create, process, store or access data in digital form"* or *"a service that allows the sharing of or any other interaction with data in digital form uploaded or created by the consumer or other users of that service"*. As mentioned in our chapter regarding these Directives, there are remedies and solutions for consumers on both situations, and it is expected that a more predictable legal framework for European traders/sellers will exist after they are implemented by Member States.

By comparison, the scope of application of the Product Liability Directive is much narrower. Product, in accordance with this Directive means *"all movables even if incorporated into another movable or into an immovable. 'Product' includes electricity"*. Therefore, and according to the interpretation shared by Commission[367] and the ECJ[368], the Product Liability Directive is applicable to a service provider using defective equipment or products of which it is not the Producer.

The problem with the Product Liability Directive's scope of application is that AI will frequently be provided as a service and not product. Furthermore, hybrids between

366 Traders for the Digital Services Directive and Sellers for the Sale of Goods Directive.
367 *See*, "Communication from the Commission: Consumer Policy Action Plan 1999-2001", European Commission, accessed August 15, 2019, http://aei.pitt.edu/6657/1/6657.pdf; "Green Paper: Liability for Defective Products", European Commission, accessed August 15, 2019, http://aei.pitt.edu/6657/1/6657.pdf. The question is also referred Commission Staff Working Document: Liability for emerging digital technologies.
368 *See*, Judgment of the ECJ of 21 December 2011, *Centre hospitalier universitaire de Besançon v. Thomas Dutrueux*; Case C-495/10, ECLI:EU:C:2011:869.

products and services will be quite frequent, creating a space for litigation in regards to where the Directive is applicable and where it is not. Software is clearly an issue in this context. Cloud-based AI software would, most likely, be considered as a service and, thus, not fall under the Product Liability Directive[369]. On the other hand, software that is essential and necessary for a determined product to work, should, probably, be considered as part of said product and, thereby, the Producer should be liable for damages caused by defects within it. In fact, and to avoid systematic inconsistency, the notion of Product should be interpreted in a compatible manner with the notion of Goods in the Sale of Goods Directive and Digital Services Directive. Therefore, a Product should be considered as including items that incorporate or are inter-connected with digital content or a digital service in such a way that the absence of that digital content or digital service would prevent the product from performing its functions. However, and for the same reason and since the concept of Good in those Directives seems to be broader than the concept of Product in the Product Liability Directive, it seems difficult to argue that digital content or digital services should be considered as products. Furthermore, while the criteria of tangibility is not included in the Product Liability Directive (only heavily inferred[370]), it is in both the Sale of Goods Directive and Digital Services Directive and, thus, a consistent interpretation between instruments seems to be reasonable.

Some theories could be called upon to reach a different conclusion. On one hand, it could be argued, with some merit, that software is not intangible at all since when it is introduced in the adequate hardware, it will cause tangible changes in world around it. If you put a videogame into a computer or a console, it will present you with tangible (and interactive) images and sounds. If a hacker is able to introduce malware (for all intents and purposes, software) into a country's power grid, it may result in the very tangible result of everyone losing power.

An additional distinction can be considered, where software that directly affects its surroundings will be considered as having tangibility and being a product, while software that can only affect its surroundings through a human being (imagine and AI-enabled research support software for case-law), will not be considered either tangible or a product. A slightly different theory argues that software instructing a computer would be a product, while software instructing a human would be a service. According to this theory, *"the*

[369] Examples of services that can incorporate AI, and potentially cause damages due to software faults are, web-based image and video editors, web-based recommendation engines or translators.
[370] Through the reference to electricity as an intangible product where the Directive is still applicable.

situation becomes more complicated in cases where the instructive information is derived from the use of software. In such cases it must be established whether the information was generated by the software itself or whether the information was generated by the individual and only made accessible through software, as in the case of, for example, a database programme" In the first situation we would a have product, while not on the later. Lastly, it also defensible that when information is manifested into a physical and tangible medium (such as a computer or smartphone) it will have material substance and be a product[371].

However, for all the academic merit of these theories, applying them to the Product Liability Directive, in straight contradiction with the Sale of Goods and Digital Services Directive, would not be coherent or in line with the legislator's thinking. If the legislator wanted to enshrine a broader definition of product, we would see some reflex of it in the definition of good in the new Directives. Furthermore, there is no sign in the reports regarding the Product Liability Directive, that this interpretation is shared by any Institutions. Moreover, some of abovementioned theories defend an interpretation of product so broad (in the name of fairness) that they empty all the meaning from the concept of service and what is generally known as a service. It would be also highly harmful for the single market and create uncertainty in manufacturers, traders, sellers and other market players to have concepts incompatible to such a degree in these Directives.

We will draw the reader's attention to the fact that, if software is contained within a physical container (a CD, pendrive or hardrive), the physical container will, without a shadow of a doubt a product. Still, the relevance of this fact is not high, as damages caused by physical containers are not a new or prevalent challenge. But, for clarity's sake, if a pendrive containing a piece of software overheats due to a manufacturing problem, the Producer shall be liable under the Product Liability Directive.

Now, since this is a minimum harmonisation directive, and quite a permissive one at that, it is possible for Member States to establish alternative regimes, where the rules regulating products under the Product Liability Directive are extended to software in general[372]. While the solution may sound attractive at first, first impressions are not always

---

371 *See, K* Alheit, *"The applicability of the EU Product Liability Directive to software"* The Comparative and International Law Journal of Southern Africa 34,2 (2001): 188-209.

372 In Greece and Lithuania, the law extends to both products and services/intangibles, thereby covering software. In Estonia the Obligations Act, considers computer software as a product and in France Courts have the same interpretation (computer software as a product). *See*, European Commission, *Evaluation of Council Directive 85/374/EEC on the approximation of laws, regulations and administrative provisions of the Member States concerning liability for defective products* (Luxembourg: Publications Office of the European Union, 2018),

right. Potentially introducing dozens of different regimes that are both incoherent between themselves and with the rest of the European legislation such as the Sale of Goods Directive and Digital Services Directive, will completely defeat the objectives of the digital single market and force European manufacturers, developers, sellers, traders, users etc. to be aware and adapt to each and every type of legislation, instead of smoothly guaranteeing cooperation and development and deployment of AI, it would certainly have a chilling effect on the market (and additionally break consumer's trust and potentially harm them). This one of those occasions where having a wrong but coherent approach is better than having some good but incoherent approaches.

However, and regarding this specific issue, the best solution would be to have one good legislative solution. Therefore, the EU should quickly move to update the Product Liability Directive, bringing its concepts to the XXI century and lining them up with those contained on the Sale of Goods and Digital Services Directive. If needed, and in the image of what was done with these Directives, different, but compatible, regimes may be enshrined for tangible Products and digital services/content.

History shows us that minimum harmonisation Directives are, generally, a very meritless (and arguably near worthless) idea. Therefore, a maximum harmonisation directive should be the way to go. While in a perfect world a Regulation could be a solution, the political will for it would probably not be there in the area of liability, and it would not be coherent with the previous legislation approved through Directives[373].

17ff.; Taivo Liivak, *"Liability of a Manufacturer of Fully Autonomous and Connected Vehicles under the Product Liability Directive"* International Comparative Jurisprudence 4,2 (2018): 178-189.

[373] *See*, "Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee on the Application of the Council Directive on the approximation of the laws, regulations, and administrative provisions of the Member States concerning liability for defective products (85/374/EEC)", European Commission, accessed August 15, 2019, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0246&from=EN; "Review Of Product Liability Rules: BEUC Position Paper", The European Consumer Organization: BEUC, accessed August 20, 2019, https://www.beuc.eu/publications/beuc-x-2017-039_csc_review_of_product_liability_rules.pdf; Janja Hojnik, *"Technology neutral EU law: digital goods within the traditional goods/services distinction"*, International Journal of Law and Information Technology 25 (2017): 63-84; Lori A. Weber, *"Bad Bytes: The Application of Strict Products Liability to Computer Software"* St. John's Law Review 66,2 (1992): 469-485.

§ 1.2.2. Defences

§ 1.2.2.1. The State-of-the-Art Defence

The Product Liability Directive was approved under a legal basis that required unanimous agreement in the Council (if approved nowadays, this requirement would probably not be applicable, see our chapter on the Legislative Procedure). Although no official record exists from it, and proving it behind a shadow of a doubt will certainly be impossible, there are reports about how the inability to agree on the wording and interpretation gave rise to the relatively unclear form in which the current State-of-the-Art Defence is written. The Commission and Council also fought regarding its inclusion or not in the final Directive.

There is an immediate limitation to the State-of-the-Art Defence: as stated above the European legal framework is not really and does not have to be really harmonised in regards to it. Member States may, at their discretion, restrict or completely suppress the protection granted by this provision to Producers. Therefore, even if we were in complete agreement with it, and if it was an example of perfect regulation on this issue – it is not! – it would still be grossly defective. If we want a single market for AI-enabled devices and the boost to the digital single market that will certainly derive from there, Producers need to have as clear a legal framework as possible. As we know, this is an area of development that carries some risk, and it is likely that companies will wary of making available their products if they do not possess full knowledge of target's legal framework.

But the limitations of the current State-of-the-Art Defence are broader than simply not being harmonised across Member States (the Directive itself suffers from this issue, in a higher or lower degree depending on the provisions). This defence may be used when "*the state of scientific and technical knowledge at the time when he put the product into circulation was not such as to enable the existence of the defect to be discovered*". In its landmark judgment in *Commission v. United Kingdom*, the ECJ clarified the scope of and possible uses of this defence under EU Law. In accordance with the judgment:

a) Compliance with industry standards and practices by itself does trigger the scope of this defence. To use it a Producer must prove compliance with the most advanced level of technical and scientific knowledge available at the time when the product was put into circulation even if it was not an industry standard.

b) *"The clause providing for the defence in question does not contemplate the state of knowledge of which the producer in question actually or subjectively was or could have been apprised, but the objective state of scientific and technical knowledge of which the producer is presumed to have been informed"*. However, the knowledge in question must have been accessible at the time. To explain the question of knowledge accessibility, the Advocate General gives the following example: *"the aspect which I have just been discussing is closely linked with the question of availability of scientific and technical knowledge, in the sense of the accessibility of the sum of knowledge at a given time to interested persons. It is undeniable that the circulation of information is affected by objective factors, such as, for example, its place of origin, the language in which it is given and the circulation of journals in which it is published. To be plain, there exist quite major differences in point of the speed in which it gets into circulation and the scale of its dissemination between a study of a researcher in the United States published in an international English-language journal and, to take an example given by the Commission, similar research carried out by an academic in Manchuria published in the local scientific journal in Chinese which does not go outside the boundaries of the region. 24. In such a situation, it would be unrealistic, I would say unreasonable, to take the view that the study published in Chinese has the same chances as the other being known to a European product manufacturer. So, I do not consider that in such a case a producer could be held liable on the ground that at the time at which he put the product into circulation the brilliant Asian researcher had discovered the defect in it. More generally, the 'state of knowledge' must be construed so as to include all data in the information circuit of the scientific community as a whole, bearing in mind, however, on the basis of a reasonableness test the actual opportunities of the information to circulate"*.

There is quite some relevance to the question of accessibility in the context of AI development. First, while certainly not on purpose, the example is very adequate to the world of AI development, as China (as explained in Chapter IV of Part I) is indeed producing a significant number of papers on AI R&D, but those papers are generally less accessible and influent than the ones coming from other countries. Second, and questions can arise in regards to knowledge produced in the European Union itself. Should an SME from the Netherlands be able to sustain this defence if the defect could be prevented by knowledge published in a small Portuguese journal? Should this be considered outside of the "state of knowledge"? Of course, the exception also covers knowledge that is kept by companies or other entities and not made available to the public, even if it does exist.

Furthermore, the level of financial efforts that can be demanded from the Producer until one must consider that there is a disproportionate effort and the defence can be used are not clear.

c) The burden of proof falls onto the Producer. That is to say, it is up to the Producer to prove that there was not scientific and/or technical knowledge available at the time the product was put into circulation that could have prevented a defect.

It seems clear, that in its current form the state-of-the art defence neither grants Producers the necessary security to Producers nor consumers the necessary protection to consume. The main issues are that it is not applied in a uniform manner across Member States and that is written in a manner that allows interpretations that can be so broad as to allow abuses or so strict as to completely drain its usefulness. Furthermore, those interpretations can vary from Member State to Member State and not every case goes to the ECJ for uniformization, creating one more source of uncertainty[374].

The problem is that removing the state-of-the-art defence outright could lead to lower investments in R&D (or, at least, shift its priorities) and slow down innovation. Companies could be tempted to only make products available in the market when there was absolute certainty that those security measures would be advanced enough to prevent defects. In an area like AI-development, this could seriously hinder innovation and create a negative cycle. Absent, real-world feedback, Producers might take more time in developing security measures. Furthermore, innovation brings with it social and competitive advantages, that could be lost. Eventually, those advantages could be enough to offset any harm caused by damages from in-development technologies. It could also create a scenario of frivolous litigation and even bring innovative firms to bankruptcy (a situation where they would be unable to provide compensation)[375].

However, consumers still need to be compensated in some manner for damages that may arise from using these technologies. Not doing so would sacrifice the high level of consumer protection build in the EU and, potentially, even hinder innovation in itself, as the concept of trustworthy AI could be in itself affected. How can you trust something that can harm you without consequences?

---

[374] As interesting as a deep analysis of the possible interpretations of the state-of-the-art defence could potentially be, this Thesis is not the most adequate place to it. It is not that if falls outside of our scope, *per se*, but there are limitations to how far we can go in each of the themes addressed. Since such a discussion would be heavy in Member State law (and our priority is EU Law) and, in itself, would probably merit and individual study we opted to leave it to the nearest opportunity. Additionally, it is our belief that, taking into account the other issues with the Product Liability Directive, any interpretation of the state-of-the-art defence would not avoid the need for new legislation. Furthermore, without, a new and deeper intervention from the ECJ or the legislator, doubts will always persist in regards to this defence.

[375] We should also underestimate the fact that the existence of the state-of-the-art defence is seen as extremely important for Producers, even if they rarely use it.

New solutions are needed to guarantee that we can achieve both these aims and, as we will see below, some options are already available[376].

§ 1.2.2.2. The Later-Defect Defence

In accordance with Article 7 (b) the Producer shall not be held liable if he manages to prove that *"having regard to the circumstances, it is probable that the defect which caused the damage did not exist at the time when the product was put into circulation by him or that this defect came into being afterwards"*. This is known as the later-defect defence.

AI enabled-devices are, generally, frequently updated. Updates are supposed to improve functionality, enhance to Product and heighten security. However, in some cases, they can also cause malfunction or even completely break the machine. If the defect is caused by a software update, it is arguable, that it did not exist at the time when the product was put into circulation and, thus, does not fall within the Product Liability Directive. In this scenario, the injured party could make use of other available types of liability, such as non-contractual liability. This does not seem to be the spirit of the law, as its objective is to not burden the Producer with liability for defects that appear in the Product after it leaves his sphere of influence, and this would not be the case. Nevertheless, clarification could be relevant to ensure that this defence is not abused in the context.

Second, and quite related to the section below regarding proving causality. It will be incredibly difficult for Producer to prove that a defect in the software powering the AI-enabled device did not exist at the time when the Product was put into circulation, if the AI is not explainable and its decisions traceable (*i.e.* if there is a black box). This may work as an incentive to foster the development of AI with these characteristics. Still, if one aspect

[376] *See*, Piotr Machnikowski, "Introduction", *European Product Liability: An Analysis of the State of the Art in the Era of New Technologies*, in Piotr Machnikowski ed., (Cambridge: Intersentia, 2017), 1-14; "Evaluation of Council Directive 85/374/EEC on the approximation of laws, regulations and administrative provisions of the Member States concerning liability for defective products", Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs (European Commission) , EY , Technopolis and VVA, accessed August 20, 2019, https://publications.europa.eu/en/publication-detail/-/publication/d4e3e1f5-526c-11e8-be1d-01aa75ed71a1/language-en; Duncan Fairgrieve and Luis González Vaqué, "Introduction", in *Product Liability in a Comparative* Perspective, Duncan Fairgriev ed (Cambridge: Cambridge University Press, 2005), 1-9; Ian Dodds-Smith and Alison Brown, *"*Recent Developments in European Product Liability*"*, in *The International Comparative Legal Guide to: Product Liability 2009* (London: Global Legal Group, 2009), 1-4; "Autonomous Systems in Aviation: Between Product Liability and Innovation", Ivo Emanuilov, accessed August 10, 2019, https://www.sesarju.eu/sites/default/files/documents/sid/2017/SIDs_2017_paper_29.pdf; "Reflection on Developmental Risk Defence Under Product Liability Law", Akinrinmade Olomu Gbade, accessed August 20, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3066191;

may lead to the abuse of this defence, the other one may empty it of any meaning. Of course, this second issue only appears if the injured party can prove causality between defect and damages which in itself is not easy[377].

The issue may be even more difficult to solve in AI-enabled devices that are capable of self-learn and learn from their surroundings. If an AI-enabled robot learns a behaviour that leads to damages the following question arises: how to assess why it learned something? Because there is a defect in its learning system or because it was poorly taught? In a manner that almost mimics the issue of nature vs. nurture in humans, – a serial killer may be a serial killer because he is a sociopath, but it may also be because he had a traumatising childhood –, an AI may equally be either "defective" at birth or be poorly taught. As with Tay's case above, there were already cases of AI's self-learning mechanisms being exploited with harmful consequences. While Article 8 of the Product Liability Directive already offers some protection to Producers in cases where someone "sabotages" AI (willingly or in a negligent manner), if there is legislation directed to this type of technology or, at least, if it is adapted to it, more specific provisions on this matter could be relevant[378].

§ 1.2.3. Proving Causality

Under the Product Liability Directive, the Producer is strictly liable for damages caused by its products. That is to say, fault is unneeded. However, the injured party still has to prove both the defect and the causal relationship between defect and damage. While proving the defect is mostly about (legitimate) expectations (see above)[379], proving the causal relationship may prove to be disproportionally difficult for the average user. As we

---

[377] Even though the ECJ has accepted that proof based on presumptions is not incompatible with the Product Liability Directive, presumptions may be an unfair burden on the Producer. *See*, Nuno Pinto Oliveira, "Toxictorts e Causalidade", in *Cadernos da Lex Medicinae - Saúde, Novas Tecnologias e Responsabilidades: Nos 30 Anos do Centro de Direito Biomédico*, coord. André Dias Pereira, Javier Barceló Doménech and Nelson Rosenvald, Vol. II (Coimbra: Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2018), 395-408.

[378] *See*, K Alheit, *"The applicability of the EU Product Liability*…188-209; Jacob Turner, *Robot Rules: Regulating Artificial*…, 91ff.

[379] The Commission considers that *"it will be necessary to provide for adequate safety levels for all types of products, taking also account of any new risks that may be posed regarding the emerging digital technologies"*. However, we do not believe that this is really needed, or feasible at all. *See*, "European Commission Staff Working Document: Liability for emerging digital technologies. Accompanying the document Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial intelligence for Europe", European Commission, access December 10, 2018, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51633.

have explained above (see our chapter on the GDPR), it is not always possible for the developer itself to fully understand an AI's decision after it employed its self-learning capabilities (especially if it continually learns in a real-world context). If the developer and/or Producer is unable to fully understand a decision by a machine, it will be equally difficult to link a defective AI-enabled decision with damages occurred. If an autonomous vehicle runs over a pedestrian in a crosswalk, one can argue that the solution is easy. We do not really need to go very deep within the AI's thinking, an autonomous vehicle is not supposed to run over people and, in normal situations, if it worked, the damages would not have occurred. However, there are equally insidious situations where proving causality appears to be more difficult. Imagine the AI-enabled drone that automatically makes deliveries in the most effective manner to ensure the fastest and most profitable delivery times. If this drone starts delaying deliveries in a certain neighbourhood, it may not be possible to understand if the AI's decision was because the average socio-economic status of the people living there is lower and thus they order cheaper products and it is, arguably, engaging in discriminatory behaviour or if, on the contrary, it is just making a defensible decision based purely on the most effective way to distribute its packages. Another example, an AI-enabled RH equipment (for argument's sake, let us suppose that it is a special machine and under the Product Liability Directive). The equipment starts discriminating against female candidates, which is detected after a few months. How can we discover if the discrimination was due failures in the initial setup of the computer program, in the original datasets (where the Producer would probably be liable) or due to the machine learning from the preferences shown by the company which is using it in the previous months and that tends to exclude female candidates at a second stage due to bias from its interviewers (where the Producer would probably not be liable).

As we argued in our chapter about the GDPR and section on information and explanation, the ability to explain an AI's decision is critical, and not only for data protection reasons. It is also essential to assess the causal relationship between defect and damage. That is investing in explainable and traceable AI is key for developments in this field.

Some Member States opted to enshrine special rules lessening the burden on the injured party, and the ECJ greenlighted this option, considering that is compatible with the Product Liability Directive. Nevertheless, and again, this introduces one more incoherent aspect in the European regulatory framework, a fact that is undesirable. Eventually, a solution such as reversing the burden of proof may be equated. But there are good reasons

to believe that this may be too heavy at this stage of development and, at least in certain areas that are more sensible, may not be desirable. Even within ML not everything is equal, and some areas where the EU wants to foster development at a higher pace or that are traditionally riskier but key to the European strategy may need sectorial approaches.

As we will see below, alternative solutions such as mandatory insurance schemes may be more adequate and provide for adequate compensation without creation of chilling effects.

§ 1.2.4. Limited Damages

Under the Product Liability Directive, compensation is limited to damages caused by death or personal injuries and damages to property with a lower threshold of 500€. Starting with the question of the lower threshold, it is understandable why it is there (to limit excessive litigation), but that issue is limited by nature due to the financial costs (and psychological burden) that comes with litigation in itself. Furthermore, with AI being more and more in our personal life, the threshold makes no sense. What the Directive could suggest, instead, is for the implementation of a quicker and simpler procedure to settle claims regarding small values.

The distinction between private and personal use may need some fine-tuning. This distinction was thought in the context of exempting damages caused to machinery, for example, in factories. However, with AI-enabled devices and working from home being more common, it may not be adequate nowadays. Imagine that the malfunction in an AI-enabled vacuum cleaner destroys the library of a small lawyer's office or the computer of self-employed web designer. In the spirit of the law, these damages should, most likely, be compensated. Including the professional and personal use dichotomy also brings with the question of dual use, that can be, and has been, quite troublesome for consumer protection law, and ideally, should be avoided. With this in mind, we should take into account that the concept of *"mainly for his own private use or consumption"* is, in fact, clear in its choice for a broad definition of private use and could, probably, stand better than some more recent alternatives enshrined in EU and national laws.

Nevertheless, under alternatives to strict liability, such as mandatory insurance schemes or compensation funds, the division appears to make even less sense and different manners of tackling the issue may be explored, such as restricting the value to be

automatically compensated by the insurance to a certain adequate amount for simpler appliances (such as simple electronics), while the person can still seek higher damages directly from the Producer.

Including compensation for other types of damages, such non-pecuniary damages may also be adequate with the growing prevalence of AI. In our example regarding the HR machine, there is clearly damages in the form of psychological and emotional harm and, arguably, even loss of earning that could be compensable. While those cannot be easily adapted to our alternative solutions such as mandatory insurance schemes, they can still be included in future regulation on this matter and compensable directly by the Producer when they do arise. This would be in line with Parliament's opinion regarding this matter expressed on the Resolution regarding civil rules on robotics.

§ 1.3. The Future Commission Guidance on the Product Liability Directive

In the fifth report on the application of the Product Liability Directive (2018), the EC concedes that certain aspects of the Directive should be clarified including the concepts of "product", "damage", "defect", and the burden of proof under it. The Commission also considers that there is a need to assess whether the Product Liability Directive will be able to address the challenges brought by emerging technologies, including AI. A Guidance on these issues as well as a report *"on the broader implications for, and potential gaps in and orientations for, the liability and safety frameworks for AI, Internet of Things and robotics"* was promised for mid-2019, but as of this writing has not been published. Importantly, the EC appears to be determined in standing by the current strict liability regime. However, it is not clear if the Institutions means it will stand by the current framework in the matters currently properly addressed by it, or also in emerging technologies.

One must keep in mind that, whatever the content of the Guidance and report, the Commission has no directive legislative power in this matter and therefore, they will be non-binding. This type of soft law instruments may be highly relevant, but it cannot extend the scope of the current Product Liability Directive and solve issues such as the lack of services' inclusion. Furthermore, if the ECJ disagrees with the CE's conclusions any interpretation by the latter institution will quickly fall. That is to say, this Guidance shall certainly be no replacement for a new legislative intervention, which is needed to avoid

regulation by a legal framework that was manifestly not though out for the current challenges and that is not even in sync with the rest of the EU's legal framework.

§ 2. Alternative Regulatory Approaches

§ 2.1. Amending the Product Liability Directive

It seems fair to state that the current Product Liability Directive, as a legal instrument, will not be able to provide the necessary answer to the challenges brought by AI.

If the objective is to contribute to the development of the single market and digital single market, by design, a minimum harmonisation Directive fails at the necessary requirements. It may be able to provide some level of protection to consumers, but even in that we would argue that it should be less effective than other instruments. Currently, the legal frameworks between Member States offer a minimum level (unsatisfactory) of protection but are not aligned in a manner that will allow manufacturers of AI-enabled products and services fluid and effortless trade across the EU. The greater impact will be felt by smaller companies and start-ups, since neither possesses highly refined in-house legal departments nor have the financial capability to hire local legal advice to provide them with the specificities of each legal framework. Remarkably, most of the EU's business fabric possesses these characteristics, and according to the preparatory documents from the European Institutions, those are the ones that the EU most wants to foster.

The Product Liability Directive is a 1985 legal instrument, lightly amended in 1999. It was not supposed to regulate technological phenomena from the current days. It is only natural that its concepts and definitions suffer from severe inadequacies such as not considering services. However, with the Sale of Goods and Digital Services Directive there is already a good conceptual basis on which to work and thus the European legislator should swiftly update the Product Liability Directive and bring in in line with the rest of the EU's legal framework. If during preparatory work for the new legal instrument it is considered that different approaches are needed for products and services, those approaches should be developed jointly, coherent and compatible. Proving causal relationship between defect and damage can be a considerable challenge for the injured party, especially for more complex AI-enabled devices, which will be more and more

common as the technology develops. To avoid leaving the users unprotected the burden of proof must be either lessened or fully reversed. However, such a decision may be negative from the Producer's point of view.

If the principle of strict liability is kept, the defences available to manufacturer's should be reviewed to take into account the new paradigm. The Later-Defect Defence should take into account updates and self-learning on the Producer's side of the equation but also the difficulties arising potential manipulation of the self-learning capabilities by malicious users. The State-of-the-Art Defence needs to be clarified and amended. It cannot stand with a wording that is unclear and provides no protection for consumers or security for manufacturers. Furthermore, Member States should not be allowed to suppress it if they wish so. Maximum harmonisation is the way to go.

If a strict liability scheme is kept, compensation should not be restricted to death, personal injury and property damages. AI will work extremely close to human beings, thereby other types of non-pecuniary damages should also be compensated. Furthermore, restricting compensation to personal use can give rise to unnecessary unclearness and even litigation due to the issue of dual use. A solution may be to suppress the distinction, as its existence arguably lost some meaning in the XXI century.

In the end, the key question should be if the strict liability regime currently in force should be kept for AI. It does not appear to be evident that doing so, at least in isolation, is the best solution.

§ 2.2. Mandatory Insurance Schemes and Compensation Funds

The European Parliament in its Civil Law on Robotics Recommendation called upon the Commission to study and consider solutions that include:

"a) establishing a compulsory insurance scheme where relevant and necessary for specific categories of robots whereby, similarly to what already happens with cars, producers, or owners of robots would be required to take out insurance cover for the damage potentially caused by their robots;

b) ensuring that a compensation fund would not only serve the purpose of guaranteeing compensation if the damage caused by a robot was not covered by insurance;

c) allowing the manufacturer, the programmer, the owner or the user to benefit from limited liability if they contribute to a compensation fund, as well as if they jointly take out insurance to guarantee compensation where damage is caused by a robot;

d) deciding whether to create a general fund for all smart autonomous robots or to create an individual fund for each and every robot category, and whether a contribution should be paid as a one-off fee when placing the robot on the market or whether periodic contributions should be paid during the lifetime of the robot".

If we exclude the question of the "specific legal status", Parliament seems to have the right idea[380]. No fault-insurance immediately kills some of the most pressing issues with regulating liability for AI, including the problems with the burden of proof, concepts of defect and when it appears, foreseeability and, in somewhat of a self-contradictory manner, specific legal status or personhood[381]. There are some areas where establishing mandatory insurance will probably not be difficult, and the EP gives a few examples (autonomous vehicles being the classical, medical and care robots being another good candidate).

Mandatory insurance schemes have their limitations, though. First, insurance needs to be available and that can only be done by the insurance providers. Second, it must be affordable, because if it is not it will have a chilling effect on development. Third, it is not applicable to every type of AI or even AI-enabled device. Should Apple have to have insurance to cover any damages that may be caused by Siri or Google by Google Now? A careful and technical study of where the solution is applicable should be undertaken and if new devices appear the law may not adapt fast enough. Fourth, some Authors consider that this may remove a deterrence and increase the number of accidents as manufacturers would be less worried about being held liable for them. Fifth, it is not clear if insurance could be used to compensate every type of damages possibly arising from defects in AI, or if some would fall back to the general liability regime.

Complementing the insurance solution, a general no-fault compensation fund or several smaller compensation funds could be created, based on contributions from players in the AI development industry. If someone was harmed in an accident caused by AI, the victim could be directly compensated by the fund, without the necessity of time-consuming and expensive litigation. The compensation scheme could be used for AI-enabled devices and services that can be traditionally considered less risk-prone than their insured counterparts. It has the additional advantage that, if contributing to the fund, a

[380] Even then, as argued above the EP only appears to have intended for the EC to explore the question for a future where more advanced AI is available.

[381] Concepts such as foreseeability and even the eventual legal status could be more relevant for contractual and non-contractual liability than for Product Liability. In our Thesis, we did not perform and analysis of either as our focus is in EU Law. However, the solutions of mandatory insurance and compensation funds can be highly relevant for all types of liability including contractual and non-contractual.

company can put a product on the market even if there is still no adequate insurance offer by insurance providers to it. In theory, a fund could be transversal and fill the loopholes that will, inevitably, be left out by the insurance scheme. Obviously, there are objections to this solution. First, there may not be political will to implement this regime, as it may be against the beliefs of less State intervention-oriented individuals. Second, companies themselves may prefer to be subject to liability claims by themselves, instead of sharing the burden. Third, deciding which companies will pay the necessary contribution and the value of the contribution can be troublesome. Fourth, mounting this fund at a European level may be difficult and resource-consuming, and the fund may not be able to compensate every type of damage. There is an advantage to this solution that could be particularly interesting for the European objectives of developing open-source software and hardware tough. It is possible to cover "open-source AI" through this fund. Users would be protected as they could still be compensated for damages and developers would also not be liable for damages that could have arisen from what is effectively non-commercial work. The question that persists is whether for-profit AI-companies should be called upon to foot the bill for open-source software, or if a State contribution is in store to cover damages caused from this type of AI.

With these objections in mind, we are confident that most of them can be overcome with careful planning and mounting of both the insurance schemes and obligations arising from it and compensation fund(s). Product Liability should still exist and be adapted to AI, but ideally would only be triggered where both these remedies could not provide an adequate response to the user. Some potential examples could be products which were not under the mandatory insure obligation, if the compensation fund was not capable of compensating certain types of damages such as psychological damages or if a limitation was put on the compensation fund and damages higher than certain value had to be compensated according to general liability rules.

§ 2.3. Why Do We Think That a Specific Legal Status is not the Solution (for now)?

From a European perspective there are significant issues with conferring a specific legal status to AI-enabled software agents or AI in general. Naturally, at the current stage of AI development we are excluding the creation of a natural person-like status for AI with

correspondent rights and obligations. What can be defended is attributing the status of legal person (or a similar formula) to AI.

First, as explained in our chapter about the legislative procedure, we actually have to have a manner of implementing this new legal status in the EU as a whole. This is our first problem, as establishing the concept of person (in our case legal person) is not as easy as it may seem. The criteria for being considered one and rights and obligations arising from it are within the Member States sovereign sphere of competences. Should we consider that this completely excludes any EU-wide legal status for AI? Not necessarily, but it certainly makes it more complicated to achieve. There are solutions within the EU's legal framework. We could consider that, even if the EU could not fully harmonise it, it would still be possible to establish that Member States have to grant a legal status to AI with a certain rules and obligations under the EU's competences regarding the single market. This, arguably, would not breach Member States competences but could create serious complications with subsidiarity. Another solution would be to use the flexibility clause under Article 352 TFEU to grant the EU any competences that it may be lacking in the area. However, this would depend on real advantages and added value being gained from this initiative or it would infringe the principle proportionality and subsidiarity. Currently we are unsure if such conditions are met. In addition, it is highly debatable that the political will to do so is there.

Now, as we have explained, if (when) we reach general AI or get nearer to that objective a specific legal status for artificial intelligence will be needed, but then we will be talking about AI that has capabilities that are akin to the ones possessed by a human. That is why studying the issue as soon as possible is relevant, because we will be debating it again, albeit in a different context, possibly in a not far away future. But, for now, is there really any advantage in a specific legal status? The answer appears to be negative.

At the current stage of development there is no good reason to make AI a subject of legal rights and duties.

The main advantage of doing so would be the fact that an AI could be considered liable for damages caused by its behaviour, removing potential liability from other agents such as the producer, seller, programmer and deployer or end-user. Of course, the other side of the coin (and immediate disadvantage) is that the AI would need to have resources to pay any compensation arising from its actions. Since AI is not expected to be paid for its "services", at least in the near future, the origins of AI-owned resources would, most

likely, be the abovementioned agents. Obviously, we are not even entertaining a perspective where the injured party would not be compensated because the AI, by design, does not have the assets to do so. One other advantage is that AI would be capable of entering into insurance contracts for itself and by itself.

There are, however, a plethora of downfalls to giving AI a specific legal status at the moment, and the closer that status is to legal personhood the worst they are. We already explained the first, which is harmonized implementation. Second, everything that we could gain from it can be equally be achieve by the solutions that we have proposed above without suffering from the same pitfalls. We can give the necessary degree of protection against liability from unpredictable AI actions to other agents (maybe not full protection but that is arguably not desirable), without significant disruption to our current legal frameworks and legal concepts. Meanwhile, our suggestions also ensure that the injured party will be compensated for the damages it suffers, and thus fosters trust in the technology. Third, ethical considerations around what would mean for a machine to be considered a legal person should also be carefully assessed before taking that step even if, in the end, we end opting for a completely functional concept of personhood. At the current stage of development of AI we are not able to understand what it will truly become and there is not enough discussion or research to take the next step with the necessary solid foundations. Fourth, legal status should come with rights and obligations, we would have to think if rights should be given to AI, what rights, when and at level of intelligence should they be given. Again, nowadays we still do not have an adequate theoretical basis for such a decision. Fifth, narrow AI is not able to achieve a human-like level of understanding of its decisions and the consequences arising from said decisions. Being held liable for actions that it cannot understand would not be compatible with the universal concepts of legal capability as recognized by modern legal systems. Gunther Teubner calls current AI-enabled software agents as *"digital slaves, but slaves with superhuman abilities"*. Still, for all their specific "superhuman abilities" the comparison is unfair to the people who suffered the true horrors of slavery, as AI does no not possess a sliver of the intellectual and moral complexity that makes up a real human being.

Even legal persons (as opposite to natural persons) such as companies, are able to "understand" their own actions through its representatives (that indeed are the ones acting in the name of the company), the situation is not comparable to AI.

As we have shown above, from the European perspective, it is possible to protect producers, sellers and consumers through the abovementioned solutions. Therefore, at the present stage, a specific legal status would offer no added value. We are confident that for other types of contractual and non-contractual liability that are currently addressed at the Member State level, the same is true. Solutions such as software agency accepted absent legal personality will probably work better for the time being. We would not be surprised to see some aspects of contractual, non-contractual liability and even criminal liability for AI being regulated at the European level in the future and, in fact, would see that decision in a positive light. A general assessment of the best solutions to doing so through the EU might be needed then. Truth be told, some aspects of contractual liability for AI, albeit a specific type, are already regulated by the EU through consumer protection legislation (see our chapter "AI-enabled devices and services and Consumer's Rights: A Short Introduction to the EU's new legislative framework") and the solutions proposed in this chapter work quite well with the current (to be applicable in the future) rules.

In the future, maybe closer than what we are all expecting, when AI is able to pass the Turing Test or is close to doing so and, thus, achieves capabilities akin to a human, the question of its legal status will have to be reviewed. That discussion will be highly challenging and probably change a great number of current conceptions in the field of legal studies. Starting the debate as soon as possible is a good decision but we should not try to implement solutions before they are needed rashness may cause a chilling effect in itself[382].

[382] *See*, Iakovina M. Kindyldi, *Smart Companies: Company* …, 28ff.; Robert van den Hoven van Genderen, "Do We Need New Legal Personhood in the Age of Robots and AI?", *in Robotics, AI and the Future of Law*, Marcelo Corrales, Mark Fenwick and Nikolaus Forgó eds., (Singapore: Springer, 2018), 15-55; Gunther Teubner, *"Digitale Rechtssubjekte? Zum privatrechtlichen Status autonomer Softwareagenten / Digital Personhood? The Status of Autonomous Software Agents in Private Law"* Ancilla Iuris 106 (2018): 106-149; Nuno Sousa e Silva, *"Direito e Robótica: uma primeira aproximação"* Revista da Ordem dos Advogados 77 (2017): 486-553; "Regulating liability for AI within the EU: Short introductory considerations", Tiago Sérgio Cabral and Francisco de Andrade, acessed October 26, 2019, https://officialblogofunio.com/2019/10/25/regulating-liability-for-ai-within-the-eu-short-introductory-considerations/

Conclusions and Policy Suggestions

There is a reason why the second chapter of this Thesis was dedicated to how laws are made in the EU. In fact, more than one reason. It would not be possible to defend European-wide harmonised AI rules, that complies with the EU's high level of standards for fundamental rights including data protection and consumer protection if those rules were not adopted in a democratic and transparent manner. For realising the goal of European AI, "trustworthy AI" if we prefer to use the designation of the European Commission, we must ensure exactly that: trust. Moreover, citizens can only trust in AI and in AI Regulation if they know how it is made and are able to participate and follow the process. AI will represent a dramatic shift in our society, and we cannot build it on an unstable foundation. By explaining and knowing EU rulemaking, we are able to counterargue against any populist argument relating to "bureaucrats in Brussels" controlling our laws on artificial intelligence.

Still, some good practices and even changes must be undertaken to ensure that we start building any type of rules relating to AI in a "trustworthy" manner themselves. Arguing for those is our second reason for including a chapter on the legislative procedure. Most legal basis that we can see for adoption of unified AI Rules in the EU would fall on the ordinary legislative procedure in any case (*e.g.* single market or transportation). However, if rules are adopted in matters where the ordinary legislative procedure is not the one enshrined in the treaties a *passerelle clause* should be deployed to change this fact. The advantages in terms of democratic legitimacy and transparency are too relevant to ignore. In all probability, Council (executive) would be subordinate to national parliaments if it was legislating this matter at a national level, it should not be able to ignore the European Parliament because legislation is being drafted at the European level. If, exceptionally, this is not possible the Council will have to make an effort to reform some of its less transparent and hardly democratic habits.

Even if the ordinary legislative procedure is adopted, transparency in trilogues is key for AI Regulation. Mistakes can be costly in this phase and should be avoided. On that note, it is important for citizens to know that they can in intervene in all manners. At the limit, and we hope it is not needed, there is always the citizens' initiative.

For now, the European Union is doing well. It is arguably late to the game but playing it as it should. After the initial jolt by Parliament (with some help from the ECON). The European Commission is taking all the right steps and from its Communications we can see that it is carefully studying how to craft a winning European strategy while, at the same, time already coordinating Member States efforts to foster investment in AI. We expect to see the efforts doubled in the "von der Leyen Commission" and with the new Multiannual Financial Framework.

Regarding specific AI-legislation, Member States and the Commission appear to favour GDPR-like rules[383]. However, we are not sure that this is the best possible approach. For what is worth, as a legal field of study, data protection is and was when the GPDR was being designed, significantly more developed than AI[384]. To avoid a drastic chilling effect and potential of opting for wrong options when legislating due to lack of knowledge, AI-specific Regulation should be mostly based on principles, with specific and more detailed rules for matters that are key or that the legislator thinks that must be implemented in a very specific manner. For example, we believe that specific rules should exist (and complement the ones in the GDPR) regarding the explainability of decisions made by AI. Transforming the assessment for Trustworthy AI contained into the Ethics Guidelines into a workable and effective test for Trustworthy AI to be implemented into legislation should also be considered, in the context of ethics by design.

Of course, principle-based legislation may be ineffective, and it is extremely important to avoid that. Therefore, mechanisms to ensure harmonized and uniform interpretation through the EU have to be designed and deployed. We believe that a European AI agency should be created and said agency should be supported by national agencies in an EDPB-like structure. Such a structure also allows for harmonized and uniform implementation of the principles contained in our future legal instruments. However, said agency will, arguably, need a stronger capacity of intervention than the one enjoyed by the EDPB and national supervisors on data protection as one would wish for a European AI agency to be capable of *"protection of public welfare through the scientific evaluation and supervision of AI products, software, systems or services"*[385].

---

383 *See*, "Next European Commission takes aim at AI", Laura Kayali, accessed September 10, 2019, https://www.politico.eu/article/ai-data-regulator-rules-next-european-commission-takes-aim/.

384 An even taking this into account the GDPR is not exactly perfect.

385 Floridi and others, compare it to the European Medicines Agency, considering that a *"a "post-release" monitoring system for AIs similar to, for example, the one available for drugs should be developed, with reporting duties for some stakeholders and easy reporting mechanisms for other users". See*, Floridi et al., *"AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations"* Mind and Machines 28,4 (2018): 689-707.

Indeed to fully enjoy the benefits that AI can bring to our society and to the EU, national and the European Agency should work with the relevant stakeholders through frequent consultations, regulatory sandboxes, innovation deals and other collaboration means to ensure that developers are able to easily develop AI that is able to comply with the European rules. Guidances and Guidelines will be also key and can be adapted in accordance with the state-of-the-art of AI development.

But there is already some highly influential legislation at the European level, that regulates AI today or will in the very near future. Let us start with the GDPR, probably the most prominent regulatory digital single market victory of the European Union. Few EU legislative acts can cause such a wide range of emotions as the GDPR. The range of emotions is wide, because depending to whom you are talking, they may describe it as the best human invention since sliced bread or the single reason why the EU is doomed to regress back to the dark ages. Emotional opinions notwithstanding, in our chapter on the GDPR, we addressed some of the most pressing issues headaches that this diploma can cause for AI developers and deployers. For what is worth, we do believe that AI and the GDPR can beautifully and harmoniously coexist, from basic principles to rights such as erasure or explanation. However, some effort will be needed to address questions that will undoubtfully arise and are, in fact, already arising. In this situation, where there are many questions and few answers, both the EDPB and national supervisory authorities should engage with stakeholders and help in the effort to build GDPR-compliant AI. Guidelines, opinions and other soft law instruments will be key but hardly sufficient. Close collaboration, with initiatives such as abovementioned regulatory sandboxes and innovation deals can also contribute to give companies the peace of mind that they need to develop compliant solutions. On this area the regulator should appear, at least initially, more in a role akin to partnership. It is important to have the power to levy fines that can act as a deterrent for future infringements, but at this stage of AI development learning to avoid infringement due to lack of knowledge or pure accident should be the priority. A future European regulator specifically in AI could work along with the current data protection regulators to boost this effort.

Of course, the EU cannot expect to achieve success in this mission to build GDPR-compliant AI-based exclusively upon the private sector's efforts. Not if it wants to be a market leader because then companies will just migrate to more "desirable" and less restrictive countries. Therefore, the EU should be fully available to support the development of AI with European values both through public funding but also by

collaborating in collecting private funding. Technologies such as explainable AI will be key to ensure that AI will both comply with our current legal framework and be trustworthy for our citizens.

There is also the question of AI that does not process personal data and, therefore, does not fall within the rules of the GDPR. The fundamental rights of European citizens should still be protected when that is the case, including through mandatory AI self-identification and explainability.

The Union should act swiftly on liability, before Member States feel that they must. If Member States start preparing their legal frameworks for liability arising from autonomous machines, it will be more difficult to achieve full harmonisation after. There are plenty of possible solutions, as we have seen above, and their nature is quite different and not always compatible. First, the Product Liability Directive needs to be amended as soon as possible. The necessity is even more pressing with the new Sale of Goods Directive and Digital Services. From concepts to scope, these instruments are not compatible with the Product Liability Directive, and they should be. There are a plethora of other problems (such as not including services), that we have analysed above and that, in our opinion, make it so that soft law is not enough to fix the Product Liability Directive. Truth be told, it is only natural that it is showing its age and the EU should give it a pleasant and dignified retirement.

An assessment should be carried out and where it is viable, to complement, the new Directive a combination of mandatory insurance schemes and compensation funds should be deployed. This will have the advantage of ensuring European citizens can still be compensated when they suffer harm caused by AI, foster confidence, lower litigation, be less burdensome on companies when compared to the general regime of strict liability and lessen the burden on not only Producer's Liability but also on other forms of liability at a national level such as contractual and non-contractual liability[386].

For now, we stand firm in our rejection of a specific legal status for AI-enabled devices and services. Not only because it is more difficult to implement at the current state (while not offering concrete advantages), but also because doing so at this stage could actually be harmful. Still, that is not to say that the issue should not be studied in theory and reassessed as AI evolves. Eventually, if we ever reach something that is similar to

[386] Or at the European level if the legislator ever decides to create a harmonized regime (maybe only for AI).

general AI it will be necessary and then, having conducted proper discussion and study beforehand, we should have no misgivings about implementing it.

Will new AI-specific legislation be needed? We believe so, and that appears to be the European leaders' thinking also. We have a right to explanation on AI that uses personal data, but not on AI that does not use (even on the former it is not as clear as it would be desirable), certain ethical and fundamental rights imperatives should be introduced into law and unified rules on testing and deployment may be needed. Nevertheless, an equilibrium has the be found between high-level principles to be complemented by soft law and eventually standards and stricter rules on matters where they are needed. If needed, the ECJ should also be ready to clarify any doubts that arise.

For both citizens and companies, we believe that it is better if legislation on this matter is harmonised in the EU and not dispersed between various legal instruments when no good reason for doing so exists. Therefore, it is our hope that, after the current preparation stage in which we currently are, the European Commission presents a proposal for regulating the relevant aspects of AI that are still in need, that manages to think and find solutions for the problem in a holistic manner. Throughout this Thesis, we stated some opinions and gave some pointers on what our opinion is regarding diverse matters for future regulation. However, at more than 100.000 words our current work is not as light as it could – and maybe should – be and it is not possible for us to analyse every single question. Some of them still need input from other areas of knowledge as philosophy and computer science, before they can be integrated into the law and therefore, even if we wanted and had the opportunity to do so, we would not be able to. What, if any, ethical principles should be integrated into AI legislation and directly into its programming is one of those challenges, that truly merits a specific analysis and we hope that, in the future, we have the opportunity to do so. How to integrate said ethical principles directly into programming is another highly interesting issue, but more for our friends and colleagues studying computer sciences[387].

In addition, even after we manage to achieve the "General Theory on the Law of Artificial Intelligence" (if we ever manage to do so), there will still be plenty of sectorial and specific challenges to tackle. How to "export" our perfectly crafted theory and rules

[387] *See,* Kevin LaGrandeur, "Emotion, Artificial Intelligence, and Ethics", *in Beyond Artificial Intelligence: The Disappearing* …, 97-109; *Tiago* Sérgio Cabral and Francisco de Andrade, acessed October 26, 2019, https://officialblogofunio.com/2019/10/25/regulating-liability-for-ai-within-the-eu-short-introductory-considerations/.

to other legal frameworks. Because AI is one of those areas where having machines that are not compliant with ethics and fundamental rights in one side of the world, may cause us serious problems in the future, even if our machines currently are. How to deal with the dramatic shift in the labour market that will come from AI replacing human workers in certain areas and creating job opportunities for which we may not have qualified professionals available. How to deal with AI in hospitals and even inside our bodies through smart pills or even augmentation? Should AI have a place in our courts? And if so, should it play a supporter role (our position) or act as judge (and maybe jury and executioner)? What about AI in war, a concept that appears to be incompatible with the principles that the EU wants to enshrine for AI development.

There is genuinely a lot left to explore… As we conclude we are absolutely confident of one thing and hopeful of another: we are absolutely confident that more explorers will keep producing high quality works on AI and AI regulation in the next years and will answer some of the questions that we were not able to, and we are hopeful that we too will have the opportunity to participate in that exploration.

References

1. Articles

"A Crise Existencial da União Europeia: Ensaio em torno da realização do projecto europeu no quadro dos desafios geopolíticos e jurídico-institucionais actuais", Tiago Sérgio Cabral and Rita de Sousa Costa, accessed in March 10, 2019, https://institutoeuropeu.eu/images/stories/documentos/Pr%C3%A9mio_Professor_Doutor_Paulo_de_Pitta_e_Cunha/A_Crise_Existencial_da_Uni%C3%A3o_Europeia.pdf;

"A dearth of legislative vetoes: Why the Council and Parliament have been reluctant to veto Commission legislation", Michael Kaeding e Kevin M. Stack, accessed December 20, 2019, http://blogs.lse.ac.uk/europpblog/2016/10/25/a-dearth-of-legislative-vetoes/;

"A Perspective on Brexit", Elaine Dewhurst, accessed February 1, 2019, https://officialblogofunio.com/2016/08/04/a-perspective-on-brexit/;

"A Universal Measure of Intelligence for Artificial Agents", Shane Legg and Marcus Hutter, accessed August 20, 2019, https://www.ijcai.org/Proceedings/05/Papers/post-0042.pdf.

"After Article 50 and Before Withdrawal: Does Constitutional Theory Require a General Election in the United Kingdom Before Brexit?"; Oliver Garner, accessed February 2, 2019 http://verfassungsblog.de/after-article-50-and-before-withdrawal-does-constitutional-theory-require-a-general-election-in-the-united-kingdom-before-brexit/;

"Ain't No Rest for the Wicked": Population, Crime, and the 2013 Government Shutdown", Ricard Gil and Mario Macis, accessed December 13, 2019 http://repec.iza.org/dp8864.pdf;

"All for one and one for all - EU consumer remedies unite", Edward Turtle and Evangelia Nitti, accessed May 30, 2019, https://www.lexology.com/library/detail.aspx?g=15152d7d-462d-4e07-b8e7-fde75d677c2e;

"An agent-based approach to consumer´s law dispute resolution" David Rua Carneiro et al., accessed May 10, 2019, https://repositorium.sdum.uminho.pt/handle/1822/19079;

"Article 50 TEU: The uses and abuses of the process of withdrawing from the EU", Steve Peers, accessed February 2, 2019, http://eulawanalysis.blogspot.pt/2014/12/article-50-teu-uses-and-abuses-of.html;

"Austria: "Cookie Walls / Paywalls" Hybrids Are Permissible?", Privacy Matters: DLA Piper's Global Privacy & Data Protection Resource, accessed July 5, 2019,

https://blogs.dlapiper.com/privacymatters/austria-cookie-walls-paywalls-hybrids-are-permissible/;

"Autonomous Systems in Aviation: Between Product Liability and Innovation", Ivo Emanuilov, accessed August 10, 2019, https://www.sesarju.eu/sites/default/files/documents/sid/2017/SIDs_2017_paper_29.pdf;

"Behind Trump's victory: Divisions by race, gender, education", Alec Tyson and Shiva Maniam, accessed March 15, 2019, http://www.pewresearch.org/fact-tank/2016/11/09/behind-trumps-victory-divisions-by-race-gender-education;

"Brexit and the European Football Market: The Consequences for the Premier League and the British Players", Tiago Sérgio Cabral and Rita de Sousa Costa, accessed in November 27, 2018, https://officialblogofunio.com/2016/07/24/brexit-and-the-european-football-market-the-consequences-for-the-premier-league-and-the-british-players/;

"Brexit and the Single Market: You say Article 50, we say Article 127?"; Tobias Lock, accessed February 3, 2017, http://verfassungsblog.de/brexit-and-the-single-market-you-say-article-50-we-say-article-127/;

"Brexit, The Supreme Court (UK) and the principle of loyalty: on the question of irrevocability of a withdrawal notice", Alessandra Silveira, accessed February 10, 2019, https://officialblogofunio.com/2017/01/26/brexit-the-supreme-court-uk-and-the-principle-of-loyalty-on-the-question-of-irrevocability-of-a-withdrawal-notice/;

"China's Social Credit Score – rating a people", Joanna Klabisch, accessed July 15, 2019, https://crossasia-repository.ub.uni-heidelberg.de/4177/1/2018_Juli_Social-Credit.pdf;

"China's Social Credit System: An Evolving Practice of Control", Rogier Creemers, accessed July 15, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3175792;

"Chronos vs. Brexit: why extending article 50 and delaying Brexit might not be a feasible solution for the EU"; Tiago Sérgio Cabral, accessed in March 3, 2019, https://officialblogofunio.com/2018/12/10/chronos-vs-brexit-why-extending-article-50-and-delaying-brexit-might-not-be-a-feasible-solution-for-the-eu/;

"COMPAS Risk Scales: Demonstrating Accuracy Equity and Predictive Parity", William Dieterich, Christina Mendoza and Tim Brennan, accessed August 15, 2019, https://assets.documentcloud.org/documents/2998391/ProPublica-Commentary-Final-070616.pdf;

"Data Protection Officer according to GDPR", André Mendes Costa, accessed July 9, 2019, https://officialblogofunio.com/2017/06/13/2014/;

"Data Protection, Data Transfers, and International Agreements: the CJEU's Opinion 1/15", Christopher Kuner, accessed July, 2019, https://verfassungsblog.de/data-protection-data-transfers-and-international-agreements-the-cjeus-opinion-115/;

"Defects of the will in software agents contracting", Francisco de Andrade, et al, accessed May 10, 2019, http://repositorium.sdum.uminho.pt/handle/1822/19096,

"Democracy, negotiation, personal ambitions and backroom deals: the moment of truth for the spitzenkandidaten", Pedro Madeira Froufe and Tiago Sérgio Cabral, accessed July 2, 2019, https://officialblogofunio.com/2019/07/02/editorial-of-july-2019/;

"Designing Theory-Driven User-Centric Explainable AI", Danding Wang et al., accessed August 20, 2019, https://www.ashrafabdul.com/pdf/xai-framework-preprint-chi2019.pdf;

"Editorial August 2016", Katarzyna Gromek-Broc, accessed in February 15, 2019, https://officialblogofunio.com/2016/08/04/editorial-august-2016/;

"Editorial of July 2016", Alessandra Silveira, accessed in March 30, 2019, https://officialblogofunio.com/2016/06/29/editorial-of-july-2016/;

"Editorial of May 2017: Europe: "Ceci c'est pas une pipe!", Pedro Madeira Froufe, "Editorial of May 2017", accessed in March 27. 2019, https://officialblogofunio.com/2017/05/01/europe-ceci-cest-pas-une-pipe/;

"European Data Protection and Freedom of Expression after Buivids: an Increasingly Significant Tension", David Erdos, accessed August 20, 2019, https://europeanlawblog.eu/2019/02/21/european-data-protection-and-freedom-of-expression-after-buivids-an-increasingly-significant-tension/;

"European Union regulations on algorithmic decision-making and a "right to explanation"", Bryce Goodman and Seth Flaxman, accessed August 17, 2019, https://arxiv.org/abs/1606.08813;

"Explainable AI Driving business value through greater understanding", Chris Oxborough et al., accessed August 20, 2019, https://www.pwc.co.uk/audit-assurance/assets/explainable-ai.pdf;

"Exploring or Exploiting? Social and Ethical Implications of Autonomous Experimentation in AI", Sarah Bird et al., accessed June 3, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2846909;

"GDPR and the Challenges of Digital Memory", Kiran Wattamwar, accessed March 15, 2019, http://sitn.hms.harvard.edu/flash/2018/gdpr-challenges-digital-memory/;

"Google v. CNIL: Is a new landmark judgment for personal data protection on the horizon?", Alessandra Silveira and Tiago Sérgio Cabral, accessed October 7, 2019, https://officialblogofunio.com/2019/09/10/editorial-of-september-2019/;

"Homeopathic Democracy: The European Power Struggle over the Spitzenkandidaten", Tiago Sérgio Cabral, accessed in March 3, 2019, https://officialblogofunio.com/2018/03/05/editorial-of-march-2018;

"How Socioeconomic and Perceived Behavioral Patterns Impact Personal Zhima Credit Score in China's Credit System (DRAFT VERSION)", Jianyin Roachell, accessed July 15, 2019, https://www.researchgate.net/publication/327867876_How_Socioeconomic_and_Perceived_Behavioral_Patterns_Impact_Personal_Zhima_Credit_Score_in_China's_Credit_System_DRAFT_VERSION;

"How the West Got China's Social Credit System Wrong", Louise Matsakislouise Matsakis, accessed July 15, 2019, https://www.wired.com/story/china-social-credit-score-system/;

"How transparency can be improved in the way EU laws are negotiated and agreed", Aidan O'Sullivan, access January 30, 2019, https://blogs.lse.ac.uk/europpblog/2016/08/18/how-transparency-can-be-improved-in-the-way-eu-laws-are-negotiated-and-agreed/;

"Humans Forget, Machines Remember: Artificial Intelligence and the Right to Be Forgotten", Eduard Fosch Villaronga, Peter Kieseberg and Tiffany Li, accessed June 9, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3018186;

"Implications of the declaration of invalidity of the Directive 2006/24 on the retention of personal data (metadata) in the EU Member States: an approach to the Tele 2 of 21 December 2016", Alessandra Silveira and Pedro Miguel Freitas, accessed June 8, 2019, https://officialblogofunio.com/2017/01/22/implications-of-the-declaration-of-invalidity-of-the-directive-200624-on-the-retention-of-personal-data-metadata-in-the-member-states-of-the-eu-an-approach-to-the--tele-2-of-21-december-20/;

"Introducing AI Explainability 360", Aleksandra Mojsilovic, accessed August 20, 2019, https://www.ibm.com/blogs/research/2019/08/ai-explainability-360/;

"Introduction to the legislative processes for EU directives matters and regulations on financial services", Slaughter and May, accessed May 4, 2019, https://www.slaughterandmay.com/what-we-do/publications-and-seminars/publications/client-publications-and-articles/i/introduction-to-the-legislative-processes-for-eu-directives-and-regulations-on-financial-services-matters.aspx;

"Is there a right to explanation' for machine learning in the GDPR?", Andrew Burt, accessed June 17, 2019, https://iapp.org/news/a/is-there-a-right-to-explanation-for-machine-learning-in-the-gdpr/.

"Is there a right to explanation' for machine learning in the GDPR?", Andrew Burt, accessed June 17, 2019, https://iapp.org/news/a/is-there-a-right-to-explanation-for-machine-learning-in-the-gdpr/;

"Latest EU directives strengthen the protection of consumers in the digital world", Wolf Theiss, accessed June 5, 2019, https://www.lexology.com/library/detail.aspx?g=689432a8-4c62-4823-a011-96d8469dd662;

"Machine Bias", Julia Angwin et al., accessed August 15, 2019, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing;

"Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm", David Silver et al., access March 15, 2019, https://arxiv.org/pdf/1712.01815.pdf;

"New perspectives on sale of consumer goods – maximum harmonization and high protection of consumers as a condition for the further development of cross-border trade in single market", Maria João Pestana de Vasconcelos, accessed May 30, 2019, https://officialblogofunio.com/2019/05/13/new-perspectives-on-sale-of-consumer-goods-maximum-harmonization-and-high-protection-of-consumers-as-a-condition-for-the-further-development-of-cross-border-trade-in-single-market/;

"New Regulation on the rules and procedures for the operation of unmanned aircraft: Part A – Its relationship with national laws", Marília Frias and Tiago Sérgio Cabral, accessed August 20, 2019, https://www.vda.pt/pt/publicacoes/insights/new-regulation-on-the-rules-and-procedures-for-the-operation-of-unmanned-aircraft-part-a-its/21300/;

"Of course you can still turn back! On the revocability of the Article 50 notification and post-truth politics"; Paolo Sandro, accessed February 2, 2019, http://verfassungsblog.de/of-course-you-can-still-turn-back-on-the-revocability-of-the-article-50-notification-and-post-truth-politics/;

"Out is out (including in relation to the Mediterranean diet…). On the Article 50 of the European Union Treaty in the light of the federative principle of European loyalty", Alessandra Silveira, accessed February 10, 2019, https://officialblogofunio.com/2016/07/07/out-is-out-including-in-relation-to-the-mediterranean-diet-on-the-article-50-of-the-european-union-treaty-in-the-light-of-the-federative-principle-of-european-loyalty/;

"Populist Constitutions – A Contradiction in Terms?", Jan-Werner Müller, accessed in February 1, 2019, http://verfassungsblog.de/populist-constitutions-a-contradiction-in-terms/;

"Processing Personal Data on the Basis of Legitimate Interests under the GDPR", Gabriela Zanfir-Fortuna and Teresa Troester-Falk, accessed June 9, 2019, https://info.nymity.com/hubfs/Landing%20Pages/Nymity%20FPF%20-%20Legitimate%20Interests%20Report/Deciphering_Legitimate_Interests_Under_the_GDPR.pdf?hsCtaTracking=9cf491f2-3ced-4f9c-9ffa-5d73a77a773e%7C7469b2ec-e91c-4887-b5db-68d407654e23

"Product Liability Through the Prism of EU Law", Kevin Rihtar, accessed September 1, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2391518;

"ProPublica's COMPAS Data Revisited", Matias Barenstein, accessed August 15, 2019, https://arxiv.org/pdf/1906.04711.pdf;

"Protecting our personal data in the 21st century: why the new EU legal framework matters", Rita de Sousa Costa and Tiago Sérgio Cabral, accessed in June 3, 2019, https://officialblogofunio.com/2016/06/20/protecting-our-personal-data-in-the-21st-century-why-the-new-eu-eu-legal-framework-matters/#_edn13;

"R (Miller) v The Secretary of State for Exiting the European Union [2016] EWHC 2768 (Admin) Realpolitik and the Revocation of an Article 50 TEU Notification to Withdraw", John Cotter", accessed February 10, 2019, https://officialblogofunio.com/2016/12/02/r-miller-v-the-secretary-of-state-for-exiting-the-european-union-2016-ewhc-2768-admin-realpolitik-and-the-revocation-of-an-article-50-teu-notification-to-withdraw/;

"Reclaiming the Truth: the role of European citizens on countering fake news", Rui Castro Vieira, accessed February 3, 2019, https://officialblogofunio.com/2017/11/29/reclaiming-the-truth-the-role-of-european-citizens-on-countering-fake-news/;

"Reflection on Developmental Risk Defence Under Product Liability Law", Akinrinmade Olomu Gbade, accessed August 20, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3066191;

"Regulating the Economic Impact of Data as Counter-Performance: From the Illegality Doctrine to the Unfair Contract Terms Directive", Philipp Hacker, accessed June 13, 2019, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3391772&downlo≤ad=yes;

"Regulating liability for AI within the EU: Short introductory considerations", Tiago Sérgio Cabral and Francisco de Andrade, acessed October 26, 2019, https://officialblogofunio.com/2019/10/25/regulating-liability-for-ai-within-the-eu-short-introductory-considerations/

"Rose-tinted glasses might prove fatal: populists and their performances after the 2017 Dutch general election", Tiago Sérgio Cabral and Rita de Sousa Costa, acesso em November 25, 2018, https://officialblogofunio.com/2017/11/07/rose-tinted-glasses-might-prove-fatal-populists-and-their-performances-after-the-2017-dutch-general-election/;

"Scenarios and Potentials of AI's Commercial Application in China", Deloitte, accessed June 20, 2019, https://www2.deloitte.com/content/dam/Deloitte/cn/Documents/innovation/deloitte-cn-innovation-ai-whitepaper-en-190118.pdf;

"Technical Response to Northpointe", Jeff Larson and Julia Angwin, accessed August 15, 2019, https://www.propublica.org/article/technical-response-to-northpointe.

"The EU General Data Protection Regulation: Powerful Tool for Data Subjects?", Enrico Peuker, accessed July 9, 2019, https://verfassungsblog.de/the-eu-general-data-protection-regulation-powerful-tool-for-data-subjects/;

"The first steps of a revolution with a set date (25 May 2018): the "new" General Data Protection regime", Pedro Madeira Froufe, accessed June 17, 2019, https://officialblogofunio.com/2018/05/25/the-first-steps-of-a-revolution-with-a-set-date-25-may-2018-the-new-general-data-protection-regime/;

"The Right to Explanation, Explained", Margot E. Kaminski, https://osf.io/preprints/lawarxiv/rgeus/download;

"The right to withdraw the notification to leave the European Union under Article 50 TEU: can we still save the marriage?", Mariana Alvim, accessed February 10, 2019, https://officialblogofunio.com/2017/07/10/the-right-to-withdraw-the-notification-to-leave-the-european-union-under-article-50-teu-can-we-still-save-the-marriage/;

"The Three Laws of Robotics in the Age of Big Data", Jack M. Balkin, accessed July 10, 201~9, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2890965;

"The Validity and Limitations of Electronic Agents in Contract Formation: A brief discussion under the Electronic Transactions Act in Australia", Adrian McCullagh, accessed April 6, 2019, https://law.uq.edu.au/files/18238/A-McCullagh_The-Validity-and-Limitations-of-Software-Agents-in-Contract-Formation.pdf;

"The voters have spoken. Brexit it is.", Catherine Barnard, accessed in February 1, 2017, https://officialblogofunio.com/2016/08/04/the-voters-have-spoken-brexit-it-is/;

"Trends shaping AI in business and main changes in the legal landscape", Ana Landeta and Felipe Debasa, accessed February 25, 2019,

https://officialblogofunio.com/2019/02/20/trends-shaping-ai-in-business-and-main-changes-in-the-legal-landscape/;

"Understanding China's AI Strategy: Clues to Chinese Strategic Thinking on Artificial Intelligence and National Security", Gregory C. Allen, accessed July 3, 2019, https://s3.amazonaws.com/files.cnas.org/documents/CNAS-Understanding-Chinas-AI-Strategy-Gregory-C.-Allen-FINAL-2.15.19.pdf?mtime=20190215104041;

"Unsupervised Learning", Peter Dayan, access February 16, 2019, http://www.gatsby.ucl.ac.uk/~Dayan/papers/dun99b.pdf;

"Unsupervised Learning", Zoubin Ghahramani, access February 15, 2019, http://mlg.eng.cam.ac.uk/pub/pdf/Gha03a.pdf;

"Who Accurately Predicted the End of the Government Shutdown?", Chris C. Martin, Emory University e Kimmo Eriksson, accessed December 11, 2018, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2609920;

"Who exactly will 'take back control'? Parliament vs executive after Brexit and the 'Great Repeal Bill'", Steve Peers, accessed February 2, 2019, http://eulawanalysis.blogspot.pt/2016/10/who-exactly-will-take-back-control.html;

A.M. Turing, *"Computing Machinery and Intelligence"* Mind 49 (1950): 433-460;

Alan Ingram, *"Geopolitical events and fascist machines: Trump, Brexit and the deterritorialisation of the West",* Political Geography 57 (2017): 91-93;

Alessandra Silveira e Pedro Madeira Froufe, *"From the Internal Market to the citizenship of rights: the protection of personal data as the jus-fundamental identity question of our times",* UNIO EU Law Journal.4,2 (2018): 3-17;

Alessandra Silveira e Pedro Miguel Freitas, *"The recent jurisprudence of the CJEU on personal data .on: implications for criminal investigation in Portugal"*, UNIO – *EU* Law Journal 3,2 (2017): 45-56;

Alexandre Veronese, Alessandra Silveira and Amanda Nunes Lopes Espiñeira Lemos*, "Artificial intelligence, Digital Single Market and the proposal of a right to fair and reasonable inferences: a legal issue between ethics and techniques"* UNIO EU Law Journal 5,2 (2019), in Press;

Anand Menon e Brigid Fowler, *"Hard or Soft? The Politics of Brexit"*, National Institute Economic Review 238 (November 2016): 4-12;

Andrew D. Selbst and Julia Powles, *"Meaningful information and the right to explanation"* International Data Privacy Law 7,4 (2017): 233-242.

Anne Rasmussen & Christine Ren, *"The consequences of concluding codecision early: trilogues and intra-institutional bargaining success"* Journal of European Public Policy 20,7 (2013): 1006-1024;

Ben Goertzel, *"Human-level artificial general intelligence and the possibility of a technological singularity A reaction to Ray Kurzweil's The Singularity Is Near, and McDermott's critique of Kurzweil"* Artificial Intelligence 171 (2007): 1161-1173;

Brent Wible, "*Filibuster vs. Supermajority Rule: From Polarization to a Consensus- and Moderation – Forcing Mechanism for Judicial Confirmations*", William & Mary Bill of Rights Journal 13,3 (2005): 923-965;

Catherine Fisk and Erwin Chemerinsk, "*The Filibuster*", Stanford Law Review 49,2 (1997): 181-254; David R. Jones, "*Explaining restraint from filibustering in the US senate*", The Journal of Legislative Studies 6,4 (2000): 53-68;

China Institute for Science and Technology Policy at Tsinghua University, C*hina AI: Development Report 2018* (Beijing; Tsinghua University, 2018);

Chris Holder et al., *"Robotics and law: Key legal and regulatory implications of the robotics age (Part I of II)"* Computer Law & Security Review 32,3 (2016): 383-402;

Christilla Roederer-Rynning and Justin Greenwood, *"The culture of trilogues"*, Journal of European Public Policy 22,8 (2015): 1148–1165;

———, *"The European Parliament as a developing legislature: coming of age in trilogues?"* Journal of European Public Policy 24,5 (2017): 735-754;

Christilla Roederer-Rynning, *"Passage to bicameralism: Lisbon's ordinary legislative procedure at ten"*, Comparative European Politics (2018): 1-17;

Corinne Bendersky, *"Resolving ideological conflicts by affirming opponents' status: The Tea Party, Obamacare and the 2013 government shutdown"* Journal of Experimental Social Psychology 53 (2014): 163-168;

David Canfield Smith, *"Programming Agents without a Programming Language"* Communications of the ACM 37,7 (1994): 55-67;

David Scott Louk e David Gamage, *"Preventing Government Shutdowns: Designing Default Rules for Budgets"* University of Colorado Law Review 86 (2015): 181-258;

David Silver et. al., *"A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play"* Science, 362,6419 (December 2018): 1140-1144;

Debbie Rabina and Anthony Cocciolo, *"US Government Websites During the 2013 Shutdown: Lessons from the Shutdown Library"* Alexandria: The Journal of National and International Library and Information Issues 25, 1-2 (2014): 21-30;

Deirdre Curtin & Päivi Leino, *"In Search of Transparency for EU Law-Making: Trialogues on the Cusp of Dawn",* Common Market Law Review 54,6 (2017): 1673–1712;

Eduard Fosch Villaronga, Peter Kieseberg and Tiffany Li, "*Humans Forget, Machines Remember: Artificial Intelligence and The Right to Be Forgotten*" Computer Law & Security Review 32,2 (2018):304-313;

Efren Díaz Díaz, *"The new European Union General Regulation on Data Protection and the legal consequences for institutions"*, Church, Communication and Culture 1,1 (2016): 206-239;

Elena Esposito, *"Algorithmic memory and the right to be forgotten on the web"* Big Data & Society (2017): 1-11;

Emmet J. Bondurant, "*The Senate Filibuster: The Politics of Obstruction*", Harvard Journal on Legislation 48 (2011):467-514;

Eric A. Posner, "Can it Happen Here?: Donald Trump and the Paradox of Populist Government", in *Chicago Public Law And Legal Theory Working Paper* (n.º 605, Chicago: University of Chicago Law School, 2017);

Floridi et al., *"AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations"* Mind and Machines 28,4 (2018): 689-707;

Francesca Martines, *"Transparency of Legislative Procedures and Access to Acts of Trilogues: Case T-540/15, De Capitani v. European Parliament"* European Papers 3,2 (2018):947-959;

Francisco Pacheco de Andrade et al., *"Software Agents and Virtual Organizations: Consent and Trust"* International Journal of Services and Operations Management 6,3 (2010): 352-361;

———, *"Contracting agents: legal personality and representation"* Artificial Intelligence and Law 15, 4 (2007): 357-373;

Francisco Pacheco de Andrade, Davide Carneiro and Paulo Novais, *"A inteligência artificial na resolução de conflitos em linha"* Scientia Ivridica: Revista de Direito Comparado Português e Brasileiro 59, 321 (2010): 1-28;

Gianclaudio Malgieri, *"Automated decision-making in the EU Member States: The right to explanation and other "suitable safeguards" in the national legislation"*, Computer Law & Security Review (2019): in Press; ";

———, *"Trade Secrets v Personal Data: a possible solution for balancing rights"*, International Data Privacy Law 6,2 (2016): 102-116;

Gianclaudio Malgieri and Giovanni Comande, *"Why a Right to Legibility of Automated Decision-Making Exists in the General Data Protection Regulation"* International Data Privacy Law 7,4 (2017): 243-265.

Gijs Jan Brandsma, *"Transparency of EU informal trilogues through public feedback in the European Parliament: promise unfulfilled",* Journal of European Public Policy (2018): 1-20;

Graça Canto Moniz, *"Finally: a coherent framework for the extraterritorial scope of EU data protection law - the end of the linguistic conundrum of Article 3(2) of the GDPR"* UNIO EU Law Journal.4,2 (2018): 105-116;

Gerard N. Magliocca, "*Reforming the Filibuster*", Northwestern University Law Review 105,1 (2011): 303-328;

Gunther Teubner, *"Digitale Rechtssubjekte? Zum privatrechtlichen Status autonomer Softwareagenten / Digital Personhood? The Status of Autonomous Software Agents in Private Law"* Ancilla Iuris 106 (2018): 106-149;

Hyacinth S. Nwana and Divine T. Ndumu, "A Perspective on Software Agents Research" *The Knowledge Engineering Review* 14,2 (1999): 125-142;

Hyacinth S. Nwana, *"Software Agents: An Overview"* The Knowledge Engineering Review, 11,3 (1996) 205-244;

Ingrid Opdebeek and Stéphanie De Somer, *"The Duty to Give Reasons in the European Legal Area a Mechanism for Transparent and Accountable Administrative Decision-Making? - A Comparison of Belgian, Dutch, French and EU Administrative Law"* Rocznik Administracji Publicznej 2 (2016): 97-148;

Jack M. Balkin, "Information Fiduciaries and the First Amendment", UC Davis Law Review 49,4 (2016): 1183-1234;

Janja Hojnik, *"Technology neutral EU law: digital goods within the traditional goods/services distinction"*, International Journal of Law and Information Technology 25 (2017): 63-84;

Jens Kober, J. Andrew Bagnell and Jan Peters, *"Reinforcement learning in robotics: A survey"* The International Journal of Robotics Research 32,11 (2013): 1238-1274;

Josh Chafet, "The Unconstitutionality of the Filibuster", Connecticut Law Review 43,4 (2011):1003-1040;

John Clarke e Janet Newman, "'*People in this country have had enough of experts': Brexit and the paradoxes of populism*", Critical Policy Studies 11,1 (2017): 101-116;

Jorge Morais de Carvalho, *"Sale of Goods and Supply of Digital Content and Digital Services – Overview of Directives 2019/770 and 2019/771"* Journal of European Consumer and Market Law 8,5 (2019): 194-201.

Joy Buolamwini and Timnit Gebru, *"Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification"* Proceedings of Machine Learning Research (Conference on Fairness, Accountability and Transparency 81 (2018): 1-15;

Julia R. Azari, *"How the News Media Helped to Nominate Trump. Political Communication"*, Political Communication 33, 4 (2016): 677-680;

K Alheit, *"The applicability of the EU Product Liability Directive to software"* The Comparative and International Law Journal of Southern Africa 34,2 (2001): 188-209;

Katharine G. Young, *"American Exceptionalism and Government Shutdowns: A Comparative Constitutional Reflection on the 2013 Lapse in Appropriations"* Boston University Law Review 94,3 (2014):991-1027;

Katja Grace et al., *"Viewpoint: When Will AI Exceed Human Performance? Evidence from AI Experts"* Journal of Artificial Intelligence Research 62 (2018): 729-754;

Laura T. Gorjanc, "*The Solution to the Filibuster Problem: Putting the Advice Back in Advice and Consent*", Case Western Reserve Law Review 54,4 (2004): 1435-1463;

Lilian Edwards and Michael Veale, *"Slave to the Algorithm? Why a 'Right to an Explanation' is Probably not the Remedy You are Looking For"*, Duke Law & Technology Review 16,1 (2017): 18-84;

Liya Ding, *"Human Knowledge in Constructing AI Systems – Neural Logic Networks Approach towards an Explainable AI"* Procedia Computer Science 126 (2018) 1561–1570;

Lori A. Weber, *"Bad Bytes: The Application of Strict Products Liability to Computer Software"* St. John's Law Review 66,2 (1992): 469-485;

Maria de Almeida Alves, *"Directive on certain aspects concerning contracts for the supply of digital content and digital services & the EU data protection legal framework: worlds colliding?",* UNIO EU Law Journal 5,2 (2019): in press.

Martin B. Gold e Dimple Gupta, "*The Constitutional Option to Change Senate Rules and Procedures: A Majoritarian Means to Overcome the Filibuster*", Harvard Journal of Law and Public Policy 28,1 (2005): 205-272;

Matthew Humerick, *"Taking AI Personally: How the E.U. Must Learn to Balance the Interests of Personal Data Privacy & Artificial Intelligence"* Santa Clara High Technology Law Journal 34,4 (2018): 393-418;

Meg Leta Ambrose, *"It's About Time: Privacy, Information Life Cycles, and the Right to be Forgotten"* Stanford Technology Law Review 16,2 (2013): 101-154;

Michael J. Gerhardt, "*The Constitutionality of the Filibuster*", Constitutional Commentary 21 (2005): 445-484; Ernest Bormann, "*The southern senators' filibuster on civil rights: Speechmaking as parliamentary stratagem*", The Southern Speech Journal 27,3 (1962): 183-194;

Michael Kaeding e Kevin M. Stack, *"Legislative Scrutiny? The Political Economy and Practice of Legislative Vetoes in the European Union"*, Journal of Common Market Studies 53, 6 (2015): 1268-1284;

Mireille Hildebrand, *"Privacy as Protection of the Incomputable Self: From Agnostic to Agonistic Machine Learning"*, Theoretical Inquiries in Law 20,1 (2019): 83-121;

Nadezhda Purtova, *"The law of everything. Broad concept of personal data and future of EU data protection law"* Law, Innovation and Technology 10,1 (2018): 40-81;

Nick Jennings and Michael Wooldridge, *"Software Agents"* IEE Review 42,1 (1996): 17-20;

Nicolas Miailhe and Cyrus Hodes, *"The Third Age of Artificial Intelligence"* Field Actions Science Reports: The Journal of Field Actions 17 (2017): 6-11;

Nuno Sousa e Silva, *"Direito e Robótica: uma primeira aproximação"* Revista da Ordem dos Advogados 77 (2017): 486-553; "Regulating liability for AI within the EU: Short introductory considerations",

Patrick Fisher, "*The filibuster and the nature of representation in the United States Senate*", Parliaments, Estates and Representation 26,1 (2006):187-195;

Paul B. Lambert, *Understanding the New European Data Protection Rules* (New York: CRC Press, 2018):197ff;

Pedro Miguel Freitas, *"The General Data Protection Regulation: an overview of the penalties' provisions from a Portuguese standpoint"* EU Law Journal.4,2 (2018): 99-104;

Raya Kardasheva, "*The Power to Delay: The European Parliament's Influence in the Consultation Procedure*", Journal of Common Market Studies 47,2 (March 2009): 385-409;

Rita de Sousa Costa, A realização do direito da protecção de dados da União Europeia através das fontes não-legislativas: dos grandes temas jurisprudenciais aos desafios do soft law, no contexto da aplicação do Regulamento Geral sobre a Protecção de Dados (Master's thesis: Universidade Católica Portuguesa, 2019);

Robert R. Hoffman, Gary Klein and Shane T. Mueller, *"Explaining Explanation For "Explainable AI"* Proceedings of the Human Factors and Ergonomics Society 2018 Annual Meeting 62,1 (2018): 197-201;

Roberto Baratta, *"Complexity of EU Law in the Domestic Implementing Process"* The Theory and Practice of Legislation 2,3 (2014): 293-308;

Ronald F. Inglehart and Pippa Norris, "*Trump, Brexit, and the Rise of Populism: Economic Have-Nots and Cultural Backlash*", Harvard Kennedy School: Faculty Research Working Paper Series (2016): 6;

Roy T. Meyers, *"Late Appropriations and Government Shutdowns: Frequency, Causes, Consequences, and Remedies"* Public Budgeting & Finance 17,3 (1997): 25-38;

Sandra Wachter, Brent Mittelstadt and Chris Russell, *"Counterfactual Explanations Without Opening the Black Box: Automated Decisions and the GDPR"* Harvard Journal of Law & Technology 31,2 (2018): 841-887;

Sandra Wachter, Brent Mittelstadt and Luciano Floridi, *"Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation"* International Data Privacy Law 7,2 (2017): 79-99;

Sara B. Hobolt, *"The Brexit vote: a divided nation, a divided continent"*, Journal of European Public Policy 23,9 (2016):1259-1277;

Sarah A. Binder, Eric D. Lawrence e Steven S. Smith, "*Tracking the Filibuster, 1917 to 1996*" American Politics Research 30,4 (2002): 406-422;

Scott R. Baker e Constantine Yannelis, *"Income Changes and Consumption: Evidence from the 2013 Federal Government Shutdown"* Review of Economic Dynamics 23 (2017): 99-124;

Seth D. Baum, Ben Goertzel and Ted G. Goertzel, *"How Long Until Human-Level AI? Results from an Expert Assessment"* Technological Forecasting & Social Change 78, 1 (2011): 185-195;

Silver et al., *"Mastering the game of Go with deep neural networks and tree search"* Nature 529 (January 2016): 484-507;

Sophie Perez Fernandes, "*O digitalismo é uma forma de humanismo – o contributo da União Europeia na formatação do humanismo digital como paradigma de vida em sociedade do século XXI*", [forthcoming];

Sophie-Charlotte Fischer, *"Artificial Intelligence: China's High-Tech Ambitions"* Center for Security Studies Analysis in Security Policy 220 (2018): 1-4;

Steven S. Smith e Hong Min Park, "*Americans' Attitudes About the Senate Filibuster*", American Politics Research 41,5 (2013):735-760;

Tadas Klimas and Jūratė Vaičiukaitė, *"The Law of Recitals in European Community Legislation"* ILSA Journal of International & Comparative Law 15,1 (2008): 1-31;

Taivo Liivak, *"Liability of a Manufacturer of Fully Autonomous and Connected Vehicles under the Product Liability Directive"* International Comparative Jurisprudence 4,2 (2018): 178-189

Tiago Sérgio Cabral and Rita de Sousa Costa, *"The European Union's existential crisis: current challenges from populism to Donald Trump"* UNIO - EU Law Journal.4,1 (2018): p. 3-15;

Tiago Sérgio Cabral, *"Robotics and AI in the European Union: opportunities and challenges"*, UNIO - EU Law Journal. 4, 2 (2018): 135-146; "

———, *"Testemunhas de Jeová e a Liberdade Religiosa no séc. XXI: Uma Análise com base no Acórdão Palau-Martinez vs. France"*, e-Pública: Revista Eletrónica de Direito Público 4,2 (2017): 196-219;

Tim Oliver, "*Fifty Shades of Brexit: Britain's EU Referendum and its Implications for Europe and Britain*", Italian Journal of International Affairs 52,1 (2017): 1-11;

Urs Gasser and Virgilio A.F. Almeida, *"A Layered Model for AI Governance"* IEEE Internet Computing 21, 6 (2017): 58-62;

Vernon Bogdanor, "Brexit, the Constitution and the Alternatives", King's Law Journal 27,3 (2016): 314-322;

Viktor Mayer-Schönberger, *delete: The Virtue of Forgetting in the Digital Age* (New Jersey: Princeton University Press, 2009): 37-58;

Yavar Bathaee, *"The Artificial Intelligence Black Box and The Failure of Intent and Causation"* Harvard Journal of Law & Technology 31,2 (2018): 890-938;

2. Books

Adrienne Mayor, *Gods and Robots: Myths, Machines, and Ancient Dreams of Technology* (Princeton: Princeton University Press, 2018);

Afonso Patrão, "Anotação ao artigo 50.º do TUE", in *Tratado de Lisboa Anotado e Comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012), 186-189;

Alessandra Silveira, "Brexit e o princípio federativo da lealdade europeia: considerações sobre o artigo 50.º do Tratado da União Europeia", in *UNIO E-book Volume I: Workshops CEDU 2016*, 331-348;

———, "Sull'esercizio delle competenze dell'Unione europea: il Parlamento portoghese e il giudizio di conformità al principio di sussidiarietà", in *The role of national parliaments in the EU integration process* (Wolters Kluwer Italia/CEDAM: Milan, 2016);

———, *Princípios de Direito da União Europeia: Doutrina e Jurisprudência* (2nd ed. Lisboa, Quid Juris, 2011);

Alina Kaczorowska, *European Union Law*, (3.ª ed., London: Routledge, 2013); 148ff.;

Amparo Albalate and Wolfgang Minker, *Semi-Supervised and Unsupervised Machine Learning: Novel Strategies* (ISTE and Wiley: Chippenham and Eastbourne, 2011);

Ana Maria Guerra Martins, "Anotação ao art.º 352.º do TFUE", in *Tratado de Lisboa Anotado e Comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., 1232-1235;

———, "Anotação ao art.º 48.º do TUE", in *Tratado de Lisboa Anotado e Comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., 176-182;

Andrew Glencross, *Why the UK Voted for Brexit: David Cameron's Great Miscalculation*, (London: Palgrave Macmillan, 2016);

Anna Wyrozumska, "Article 50 [Voluntary Withdrawal from the Union], in *Treaty on European Union (TEU): a Commentary,* eds. Hermann-Josef Blanke e Stelio Mangiameli, (Heidelberg: Springer, 2013), 1385-1418;

António Gameiro, *O Papel dos Parlamentos Nacionais na União Europeia* (Coimbra: Coimbra Editora, 2011);

Arlindo Oliveira*, Inteligência Artificial* (Lisboa: Fundação Francisco Manuel dos Santos, 2019);

B.W. Schermer, *Software agents, surveillance, and the right to privacy: a legislative framework for agent-enabled surveillance* (Leiden: Leiden University Press, 2007): 17-35;

Ben Goertzel and Cassio Pennachin eds., *Artificial General Intelligence* (Berlin: Springer, 2007);

Benjamin Angel Chang, "AI and US-China Relations", in *AI, China, Russia, and the Global Order: Technological, Political, Global, and Creative Perspectives*, in Nicholas D. Wright ed (Virginia: United States Department of Defense, 2018),107-111;

César Cortes e Paulo Rangel, "Anotação ao art.º 294.º do TFUE", in *Tratado de Lisboa Anotado e Comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012), 1050-1059;

David Danks, "Learning", *The Cambridge Handbook of Artificial Intelligence*, in Keith Frankish and William M. Ramsey eds. (Cambridge: Cambridge University Press), 151-167;

Dominic Barton et al., "*Artificial Intelligence: Implications for China"* (New York: McKinsey Global Institute, 2017);

Dulce Lopes and Paula Veiga, "Anotação ao artigo 11.º do TUE", in *Tratado de Lisboa Anotado e Comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords. (Coimbra: Almedina, 2012), 54-57;

Duncan Fairgrieve and Luis González Vaqué, "Introduction", in *Product Liability in a Comparative Perspective*, Duncan Fairgriev ed. (Cambridge: Cambridge University Press, 2005);

Eileen Denza, "Article 48 [Treaty Revision Procedures]", in *The Treaty on European Union (TEU): a Commentary,* Hermann-Josef Blanke and Stelio Mangiameli eds. (Heidelberg: Springer, 2013), 1331-1355;

Elaine Rich, Kevin Knight and Shivashankar B. Nair's, *Artificial Intelligence*, (3rd ed., New York: McGraw-Hill, 2009);

Francisco Pacheco de Andrade, "Da Contratação Electrónica – Em Particular da Contratação Electrónica Inter-Sistémica Inteligente" (PhD Diss., University of Minho, 2008);

Francisco de Pacheco de Andrade et al., "Agents, Trust and Contracts", *in Information Communication Technology Law, Protection and Access Rights: Global Approaches and Issues*, Irene Maria Portela and Maria Manuela Cruz-Cunha eds. (Hershey: IGI Global, 2010): p. 188-199;

———, "Software Agents as Legal Persons", in *Virtual Enterprises and Collaborative Networks: IFIP 18th World Computer Congress TC5 / WG5.5 - 5th Working Conference on Virtual Enterprises 22–27 August 2004 Toulouse, France*, Luis M. Camarinha-Matos ed., (Heidelberg: Springer, 2004), 123-132;

Francisco Pacheco de Andrade and Teresa Coelho Moreira, "Personal Data and Surveillance: The Danger of the "Homo Conectus"", *Intelligent Environments 2016: Workshop Proceedings of the 12th International Conference on Intelligent Environments*, in Paulo Novais and Shin'ichi Konomi eds. (Amsterdam: IOS Press, 2016), 115-124;

Francisco Pacheco de Andrade, José Neves and Paulo Novais, "Software Agents and Contracts", *in Encyclopedia of Networked and Virtual Organizations,* Goran Putnik and Manual Cunha ed. (Pennsylvania: Idea Group Reference, 2008), 1-7;

Francisco Pacheco de Andrade, Pedro Miguel Freitas and Teresa Coelho Moreira, "Data Protection and Biometric Data: European Union Legislation", *in Biometric Security and Privacy: Opportunities & Challenges in The Big Data Era*, Richard Jiant et al. ed., (Switzerland: Springer: 2017), 413-421;

Gregory Koger, *Filibustering: A Political History of Obstruction in the House and Senate* (Chicago: The University of Chicago Press, 2010);

Iakovina M. Kindyldi, *Smart Companies: Company & Board Members Liability in the Age of AI* (Master's thesis: Tilburg University, 2018);

Ian Dodds-Smith and Alison Brown, *"*Recent Developments in European Product Liability*"*, in *The International Comparative Legal Guide to: Product Liability 2009* (London: Global Legal Group, 2009);

Inês Morgado, "Anotação ao art.º 293.º do TFUE", in *Tratado de Lisboa Anotado e Comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords. (Coimbra: Almedina, 2012), 1048-1049;

Ira Rubinstein, "The Future of Self-Regulation is Co-Regulation", *in The Cambridge Handbook of Consumer Privacy, From Cambridge University Press*, Evan Selinger, Jules Polonetsky and Omer Tete eds., (Cambridge: Cambridge University Press, 2018);

Ivan M. Havel, "On the Way to Intelligence Singularity", *in Beyond Artificial Intelligence: Contemplations, Expectations, Applications*, Jozef Kelemen, Jan Romportl, and Eva Zackova eds., (Berlin: Springer, 2013), 3-26.

Jacob Turner, *Robot Rules: Regulating Artificial Intelligence* (Cham: Palgrave Macmillan, 2019),

Jeffrey Ding, "The Interests Behind China's AI Dream", in *AI, China, Russia, and the Global Order: Technological, Political, Global, and Creative Perspectives*, in Nicholas D. Wright ed., (Virginia: United States Department of Defense, 2018), 37-41;

Jeffrey M. Bradshaw "An Introduction to Software Agents", *in* Jeffrey M. Bradshaw ed., *Software Agents* (Cambridge: MIT Press, 1997), 3-46;

Jerry Kaplan, *Artificial Intelligence: What Everyone Needs to Know*, (Oxford: Oxford University Press, 2016);

João Mota de Campos, João Luís Mota de Campos e António Pinto Ferreira, *Manual de direito europeu: o sistema institucional. A ordem jurídica e o ordenamento económico da União Europeia*, (7.ª ed., Coimbra: Coimbra Editora, 2014);

Jordi Bieger, Kristinn R. Thórisson and Deon Garrett, "Raising AI: Tutoring Matters", *in Artificial General Intelligence: 7th International Conference, AGI 2014, Quebec City, QC, Canada, August 1-4, 2014*, Ben Goertzel, Laurent Orseau and Javier Snaider eds, (New York: Springer, 2014), 1-10;

Justin Greenwood and Christilla Roederer-Rynning, "Taming Trilogues: "The EU's Law-Making Process in a Comparative Perspective", in *The European Parliament in times of EU crisis: Dynamics and Transformations,* Olivier Costa ed., (Maastricht: Palgrave Macmillan, 2019), 121-141;

Kevin LaGrandeur, "Emotion, Artificial Intelligence, and Ethics", *in n Beyond Artificial Intelligence: The Disappearing Human-Machine Divide*, Jan Romportl, Eva Zackova and Jozef Kelemen eds. (Cham: Springer, 2015), 97-109;

Lior Rokach and Oded Maimon, "Supervised Learning", in *Data Mining and Knowledge Discovery Handbook*, Lior Rokach and Oded Maimon eds., (2.º Ed, New York: Springer, 2010);

Llio Humphreys et al., "Mapping Recitals to Normative Provisions in EU Legislation to Assist Legal Interpretation", *in Legal Knowledge and Information Systems* Vol. 279 Antonino Rotolo ed., (Amsterdam: IOS Press, 2015), 41-49;

Lorna Woods and Philippa Watson, *EU Law* (11th ed., Oxford: Oxford University Press, 2012);

Luis Jimena Quesada, "The Revision Procedures of the Treaty", in *The European Union after Lisbon: Constitutional Basis, Economic Order and External Action*, Hermann-Josef Blanke e Stelio Mangiameli eds. (Heidelberg: Springer, 2012), 323-342;

Luís Miguel País Antunes, "Anotação ao artigo 7.º do TUE", in *Tratado de Lisboa Anotado e Comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012), 43-46;

M. Emre Celebi and Kemal Aydin (eds.), *Unsupervised Learning Algorithms*, (London: Springer, 2016);

Manuel Lopes Aleixo, "Anotação ao art.º 296.º do TFUE", in *Tratado de Lisboa Anotado e Comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012), 1060-1062;

Manuel Lopes Porto, "Anotação ao art.º 311.º do TFUE", in *Tratado de Lisboa Anotado e Comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012), 1099-1102;

Margarida Afonso, "Anotação ao art.º 218.º do TFUE", in *Tratado de Lisboa Anotado e Comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012), 832-837;

Max Craglia (Ed.), *Artificial Intelligence: A European Perspective* (Luxembourg: Publications Office of the European Union, 2018): 45-51;

Max Tegmark, *Life 3.0: Being Human in the Age of Artificial Intelligence* (New York: Alfred A. Knopf, 2017);

Michael Szollosy, "Why Are We Afraid of Robots? The Role of Projection in the Popular Conception of Robots", *in Beyond Artificial Intelligence: The Disappearing Human-Machine Divide*, Jan Romportl, Eva Zackova and Jozef Kelemen eds. (Cham: Springer, 2015), 121-131

———, "Anotação ao art.º 289.º do TFUE", in *Tratado de Lisboa anotado e comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012), 1034-1038;

Miguel Gorjão-Henriques, *Direito da União: História, Direito, Cidadania, Mercado Interno e Concorrênncia* (9th ed., Coimbra: Almedina, 2019);

Nils J. Nilsson, *The Quest for Artificial Intelligence: A History of Ideas and Achievements* (Cambridge: Cambridge University Press, 2010),

Nizan Geslevich Packin and Yafit Lev Aretz, "Learning Algorithms and Discrimination", *in Research Handbook on the Law of Artificial Intelligence*, Woodrow Barfield and Ugo Pagallo eds. (Cheltenham: Edgar Elgar Publishing, 2018), 88-113;

Nuno Pinto Oliveira, "Toxictorts e Causalidade", in *Cadernos da Lex Medicinae - Saúde, Novas Tecnologias e Responsabilidades: Nos 30 Anos do Centro de Direito Biomédico*, André Dias Pereira, Javier Barceló Doménech and Nelson Rosenvald coord., Vol. II (Coimbra: Instituto Jurídico da Faculdade de Direito da Universidade de Coimbra, 2018), 395-408;

Olivier Chapelle, Bernhard Schölkopf, and Alexander Zien (eds.), *Semi-Supervised Learning* (Massachusetts: MIT Press, 2006);

Patrícia Calvão Teles, "Anotação ao artigo 24.º do TFUE", in *Tratado de Lisboa Anotado e Comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012),

Paul Craig and Gráinne de Búrca, EU Law: Text, Cases and Materials, (6.ª ed., Oxford: Oxford University Press, 2015);

Paul Voigt and Axel von dem Bussche, *The EU General Data Protection Regulation (GDPR): A Practical Guide* (Cham: Springer: 2018);

Pedro Domingos, *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World* (New York: Basic Books, 2015);

Pedro Infante Mota, "Acordos Mega-Regionais", *in União Europeia – Reforma ou Declínio*, Eduardo Paz Ferreira coord. (Lisboa:Vega, 2016);

Pedro Madeira Froufe, "O insustentável peso democrático do populismo: deambulações em torno da União Europeia, de olhos postos em Donald Trump", in *UNIO E-book Volume I: Workshops CEDU 2016*, Alessandra Silveira coord. (Braga: CEDU, 2016): 301-311;

Pedro Miguel Freitas, Francisco Andrade, and Paulo Novais, "Criminal Liability of Autonomous Agents: From the Unthinkable to the Plausible", in *AI Approaches to the Complexity of Legal Systems: AICOL 2013 International Workshops, AICOL-IV@IVR Belo Horizonte, Brazil, July 21–27, 2013 and AICOL-V@SINTELNET-JURIX, Bologna, Italy, December 11, 2013, Revised Selected Papers*, Pompeu Casanovas et al (Heidelberg: Springer, 2014): p. 145-156;

Pierre Mathijsen and Peter Dyrberg, *Mathijsen's Guide to European Union Law* (11th ed., London: Sweet & Maxwell Ltd, 2013);

Piotr Machnikowski, "Introduction", *European Product Liability: An Analysis of the State of the Art in the Era of New Technologies*, in Piotr Machnikowski ed., (Cambridge: Intersentia, 2017);

Ray Kurzweil, *The Age of Intelligent Machines* (Cambridge, MA: MIT Press, 1992);

Ricardo Bayão Horta, "Anotação ao artigo 49.º do TUE", in *Tratado de Lisboa anotado e comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012), 183-185;

Richard E. Neapolitan and Xia Jiang *Artificial Intelligence: With an Introduction to Machine Learning,* (CRC Press: Florida, 2018);

Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction"* (Massachusetts, MIT Press, 2018);

Robert Schiitze, *European constitutional law*, (Cambridge: Cambridge University Press, 2012);

Robert van den Hoven van Genderen, "Do We Need New Legal Personhood in the Age of Robots and AI?", *in Robotics, AI and the Future of Law*, Marcelo Corrales, Mark Fenwick and Nikolaus Forgó eds., (Singapore: Springer, 2018);

Ronald Leenes and Silvia De Conca, "Artificial Intelligence and Privacy- AI Enters the House Through the Cloud", *in in Research Handbook on the Law of Artificial Intelligence*, Woodrow Barfield and Ugo Pagallo eds. (Cheltenham: Edgar Elgar Publishing, 2018), 280-306.

Rudolf Hrbek, "The Role of National Parliaments in the EU", in *The European Union after Lisbon: Constitutional Basis, Economic Order and External Action*, Hermann-Josef Blanke e Stelio Mangiameli eds. (Heidelberg: Springer, 2012), 129-157;

Rui Manuel Moura Ramos, "Anotação aos art.os 18.º a 23.º do TFUE", in *Tratado de Lisboa anotado e comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012), 258-263;

———, "Anotação aos artº 25 do TFUE, in *Tratado de Lisboa anotado e comentado,* Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012), 268;

SAE International, J3016 – Jun 2018 Standard - Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles (Pennsylvania/Michigan: SAE International, 2018);

Samantha Hoffman, "Managing the State: Social Credit, Surveillance and the CCP's Plan for China", in *AI, China, Russia, and the Global Order: Technological, Political, Global, and Creative Perspectives*, in Nicholas D. Wright ed (Virginia: United States Department of Defense, 2018), 42-47;

Shazeda Ahmed, "Credit Cities and the Limits of the Social Credit System", in *AI, China, Russia, and the Global Order: Technological, Political, Global, and Creative Perspectives*, in Nicholas D. Wright ed (Virginia: United States Department of Defense, 2018), 48-54

Stelio Mangiameli and Katharina Pabel, "Article 7 [The Principles of the Federal Coercion]", in *The Treaty on European Union (TEU): a Commentary,* Hermann-Josef Blanke and Stelio Mangiameli eds. (Heidelberg: Springer, 2013), 349-373;

Stuart Armstrong, *Smarter than us: the rise of machine intelligence* (Berkeley: Machine Intelligence Research Institute, 2014): chapter 3;

Stuart J. Russell and Peter Norvig, *Artificial Intelligence: A Modern Approach"*, (3.ª Ed, Essex: Pearson Education Limited, 2016);

Susanna Fortunato, "Article 49 [Accession to the Union]", in *The Treaty on European Union (TEU): a Commentary,*. Hermann-Josef Blanke and Stelio Mangiameli eds. (Heidelberg: Springer, 2013), 1357-1383;

Tiago Sérgio Cabral, "Democracia, legitimidade e competência legislativa na União Europeia", E-book UNIO/CONPEDI Vol. 2: Interconstitucionalidade: Democracia e Cidadania de Direitos na Sociedade Mundial - Atualização e Perspectivas, *in* Alessandra Silveira coord. (Braga: CEDU, 2018), 265-292;

Vincenzo Guizzi, *manuale di diritto e politica dell' Unione Europea* (4th ed., Naples: Editoriale Scientifica: 2015);

Vital Moreira, "Anotação ao art.º 223.º do TFUE", in *Tratado de Lisboa anotado e comentado,*. Manuel Lopes Porto and Gonçalo Anastácio coords., (Coimbra: Almedina, 2012), 847-850;

Wim Voermans, Maarten Stremler and Paul Cliteur, *Constitutional Preambles: A Comparative Analysis* (Northampton: Edward Elgar Publishing, 2017);

Wolfgang Ertel, *Introduction to Artificial Intelligence,* 2nd ed. (Cham: Springer: 2011);

Yujia He, *How China is preparing for an AI-powered Future* (Washington D.C.: Woodrow Wilson International Center for Scholars, 2017).

3. Government Publications

"(Third) Report From The Commission To The European Parliament, The Council And The European Economic And Social Committee on the Application of the Council Directive on the approximation of the laws, regulations, and administrative provisions of the Member States concerning liability for defective products (85/374/EEC)", European Commission, accessed 15 August, 2019, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0246&from=en;

"A Definition of AI: Main Capabilities and Disciplines", High-Level Expert Group on Artificial Intelligence, accessed April 10, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56341;

"A Practical Guide to Controller-Processor Contracts", Data Protection Commission, accessed June 10, 2019, https://www.dataprotection.ie/sites/default/files/uploads/2019-06/190624%20Practical%20Guide%20to%20Controller-Processor%20Contracts.pdf;

"AI Portugal 2030: An innovation and growth strategy to foster Artificial Intelligence in Portugal in the European context", INCoDe.2030, accessed July 10, 2019, https://www.incode2030.gov.pt/sites/default/files/incode_aiportugal2030_june19.pdf;

"AI Sector Deal", British Government, accessed June 14, 2019, https://www.gov.uk/government/publications/artificial-intelligence-sector-deal/ai-sector-deal;

"AI4Belgium", AI4Belgium coalition, July 11, 2019, https://www.ai4belgium.be/wp-content/uploads/2019/04/report_en.pdf;

"Amazon Machine Learning: Developer Guide", Amazon, accessed August 15, 2019, https://docs.aws.amazon.com/machine-learning/latest/dg/machinelearning-dg.pdf#types-of-ml-models;

"Anonymisation: managing data protection risk code of practice", ICO, accessed September 6, 2019, https://ico.org.uk/media/1061/anonymisation-code.pdf;

"Artificial intelligence and privacy", Datatilsynet, accessed May 15, 2019, https://www.datatilsynet.no/globalassets/global/english/ai-and-privacy.pdf;

"Artificial Intelligence Ecosystem", Task Force on Artificial Intelligence of the Agency for Digital Italy, accessed June 11, 2019, https://ia.italia.it/en/ai-in-italy/;

"Artificial Intelligence Strategy", Germany's Federal Government, accessed June 10, 2019, https://www.ki-strategie-deutschland.de/home.html?file=files/downloads/Nationale_KI-Strategie_engl.pdf;

"Artificial Intelligence Technology Strategy", Strategic Council for AI Technology, accessed August 13, 2019, https://www.nedo.go.jp/content/100865202.pdf.

"Automatic Driving Systems: A Vision for Safety", US Department of Transportation, access April 5, 2018, https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/13069a-ads2.0_090617_v9a_tag.pdf;

"Commission Implementing Decision 2019/419/EU of 23 January 2019 pursuant to Regulation (EU) 2016/679 of the European Parliament and of the Council on the adequate protection

of personal data by Japan under the Act on the Protection of Personal Information", European Commission, accessed February 12, 2019, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019D0419&from=EN;

"Communication from the Commission to the European Parliament and the Council on guidance for better transposition and application of Directive 2004/38/EC on the right of citizens of the Union and their family members to move and reside freely within the territory of the Member States", European Commission, accessed May 3, 2019, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52009DC0313&from=EN;

"Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions - Coordinated Plan on Artificial Intelligence", European Commission, January 15, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56018;

"Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions - Coordinated Plan on Artificial", European Commission – Annex", January 15, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=56017.

"Communication from the Commission: Consumer Policy Action Plan 1999-2001", European Commission, accessed August 15, 2019, http://aei.pitt.edu/6657/1/6657.pdf;

"Complexity of EU law in the domestic implementing process", Roberto Baratta, accessed May 2, 2019, http://ec.europa.eu/dgs/legal_service/seminars/20140703_baratta_speech.pdf;

"Decision of the European Ombudsman setting out proposals following her strategic inquiry OI/8/2015/JAS concerning the transparency of Trilogues", European Ombudsman, accessed January, 20, 2019, https://www.ombudsman.europa.eu/en/decision/en/69206;

"Dutch Digitalisation Strategy: Getting the Netherlands ready for the digital future", Ministry of Economic Affairs and Climate Policy, accessed June 11, 2019, https://www.government.nl/binaries/government/documents/reports/2018/06/01/dutch-digitalisation-strategy/Dutch+Digitalisation+strategy+def.pdf;

"Estonia accelerates artificial intelligence development", Estonian Government, accessed July 10, 2019, https://e-estonia.com/estonia-accelerates-artificial-intelligence/;

"Estonia's AI Strategy", Estonian Government, accessed July 10, 2019 https://www.riigikantselei.ee/sites/default/files/riigikantselei/strateegiaburoo/eesti_tehisintellekti_kasutuselevotu_eksperdiruhma_aruanne.pdf;

"Estrategia Española de I+D+I en Inteligencia Artificial", Secretaría General de Coordinación de Política Científica del Ministerio de Ciencia, Innovación y Universidades and Grupo de

Trabajo en Inteligencia Artificial GTIA, accessed June 12, 2019, http://www.ciencia.gob.es/stfls/MICINN/Ciencia/Ficheros/Estrategia_Inteligencia_Artificial_IDI.pdf;

"European Civil Law Rules in Robotics" Nathalie Nevejans (requested by the European Parliament's Committee on Legal Affairs), access January 10, 2019, http://www.europarl.europa.eu/RegData/etudes/STUD/2016/571379/IPOL_STU(2016)571379_EN.pdf;

"European Commission Staff Working Document: Liability for emerging digital technologies. Accompanying the document Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions: Artificial intelligence for Europe", European Commission, access December 10, 2018, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=51633;

"European Parliament resolution of 11 March 2014 on public access to documents (Rule 104(7)) for the years 2011-2013", European Parliament, access January 20, 2019, https://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P7-TA-2014-0203+0+DOC+XML+V0//EN&language=EN;

"European Parliament resolution of 14 September 2017 on transparency, accountability and integrity in the EU institutions", European Parliament, access January 23, 2019, https://www.europarl.europa.eu/sides/getDoc.do?type=TA&language=EN&reference=P8-TA-2017-0358;

"European Parliament resolution of 28 April 2016 on public access to documents (Rule 116(7)) for the years 2014-2015", European Parliament, access January 23, 2019, https://www.europarl.europa.eu/sides/getDoc.do?type=TA&reference=P8-TA-2016-0202&language=EN#def_1_10;

"Evaluation of Council Directive 85/374/EEC on the approximation of laws, regulations and administrative provisions of the Member States concerning liability for defective products", Directorate-General for Internal Market, Industry, Entrepreneurship and SMEs (European Commission), EY, Technopolis and VVA, accessed August 20, 2019, https://publications.europa.eu/en/publication-detail/-/publication/d4e3e1f5-526c-11e8-be1d-01aa75ed71a1/language-en;

"Executive Order 13859 of February 11, 2019: Maintaining American Leadership in Artificial Intelligence", Donald J. Trump, accessed June 10, 2019, https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/.

"Executive Order on Securing the Information and Communications Technology and Services Supply Chain", Donald J. Trump, accessed July 30, 2019, https://www.whitehouse.gov/presidential-actions/executive-order-securing-information-communications-technology-services-supply-chain/;

"Finland's Age of Artificial Intelligence: Turning Finland into a leading country in the application of artificial intelligence – Objective and recommendations for measures", Finish Ministry of Economic Affairs and Employment, accessed 10 June 2019, http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160391/TEMrap_47_2017_verkkojulkaisu.pdf?sequence=1&isAllowed=y;

"For a Meaningful Artificial Intelligence: Towards a French and European Strategy", Cédric Villani et al., accessed May 7, 2019, https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf;

"GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do;

"Government AI Readiness Index 2017", Oxford Insights, accessed July 5, 2019, https://www.oxfordinsights.com/government-ai-readiness-index;

"Green Paper: Liability for Defective Products", European Commission, accessed August 15, 2019, http://aei.pitt.edu/6657/1/6657.pdf

"Growing the artificial intelligence industry in the UK, Dame Wendy Hall and Jérôme Pesenti, accessed June 5, 2019, https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf;

"Growing the artificial intelligence industry in the UK", accessed June 14, 2019, https://www.gov.uk/government/publications/growing-the-artificial-intelligence-industry-in-the-uk;

"Guía de Privacidad desde el Diseño", AEPD, accessed 19 October 2019, https://www.aepd.es/media/guias/guia-privacidad-desde-diseno.pdf.;

"Guidelines 2/2019 on the processing of personal data under Article 6(1)(b) GDPR in the context of the provision of online services to data subjects", EDPB, accessed October 27, 2019, https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines-art_6-1-b-adopted_after_public_consultation_en.pdf;

"Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679", WP29, accessed May 5, 2019,

https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=49826

"Guidelines on consent under Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=51030;

"Guidelines on the right to data portability", WP29, accessed May 20, 2019, https://ec.europa.eu/newsroom/document.cfm?doc_id=44099,

"Guidelines on transparency under Regulation 2016/679", WP29, accessed 15 June 2019, https://ec.europa.eu/newsroom/article29/document.cfm?action=display&doc_id=51025;

"IESE Cities in Motion Index", IESE Business School, accessed July 10, 2019, https://media.iese.edu/research/pdfs/ST-0509-E.pdf;

"Intelligenza Artificiale: al via il gruppo di lavoro per una strategia nazionale", Consiglio Nazionale delle Ricerche, accessed June 10, 2019, https://www.cnr.it/it/news/8593/intelligenza-artificiale-al-via-il-gruppo-di-lavoro-per-una-strategia-nazionale;

"Italian Observatory on Artificial Intelligence", Task Force on Artificial Intelligence of the Agency for Digital Italy, accessed June 11, 2019, https://ia.italia.it/en/ai-observatory/;

"ITUC Global Rights Index from the 2018", International Trade Union Confederation, accessed July 5, 2019, https://www.ituc-csi.org/IMG/pdf/ituc-global-rights-index-2018-en-final-2.pdf;

"Machine learning: the power and promise of computers that learn by example", The Royal Society, accessed June 14, 2019, https://royalsociety.org/~/media/policy/projects/machine-learning/publications/machine-learning-report.pdf;

"Opening Statement in the European Parliament Plenary Session by Ursula von der Leyen, Candidate for President of the European Commission", European Commission, accessed July 17, 2019, http://europa.eu/rapid/press-release_SPEECH-19-4230_en.htm.;

"Opinion 05/2014 on Anonymisation Techniques", WP29, accessed September 5, 2019, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf;

"Opinion 06/2014 on the notion of legitimate interests of the data controller under Article 7 of Directive 95/46/EC", WP29, accessed June 8, 2019, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp217_en.pdf;

"Opinion 1/2010 on the concepts of "controller" and "processor", WP29, accessed June 10, 2019, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2010/wp169_en.pdf; "

"Opinion 4/2007 on the concept of personal data ", WP29, accessed September 6, 2019, https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf;

"Opinion 4/2017 on the Proposal for a Directive on certain aspects concerning contracts for the supply of digital content", EDPS, accessed June 8, 2019, https://edps.europa.eu/sites/edp/files/publication/17-03-14_opinion_digital_content_en.pdf.;

"Parecer n.º 20/2018 da CNPD, relativo à Proposta de Lei 120/XVIII", accessed in June 2, , 2019, 34v, http://app.parlamento.pt/webutils/docs/doc.pdf?path=6148523063446f764c324679626d56304c334e706447567a4c31684a53556c4d5a5763765130394e4c7a464451554e45544563765247396a6457316c626e527663306c7561574e7059585270646d46446232317063334e686279396a5a57593359544d794f4330325a44526c4c54526c4e546b74596a41304e4331694e54426d4f5449314d6a64684d7a45756347526d&fich=cef7a328-6d4e-4e59-b044-b50f92527a31.pdf&Inline=true;

"Report From The Commission To The European Parliament, The Council And The European Economic And Social Committee: Fourth report on the application of Council Directive 85/374/EEC of 25 July 1985 on the approximation of the laws, regulations and administrative provisions of the Member States concerning liability for defective products amended by Directive 1999/34/EC of the European Parliament and of the Council of 10 May 1999", accessed 15 August, 2019, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52011DC0547&from=EN

"Report from the Commission to the European Parliament, the Council and the European Economic and Social Committee on the Application of the Council Directive on the approximation of the laws, regulations, and administrative provisions of the Member States concerning liability for defective products (85/374/EEC)", European Commission, accessed August 15, 2019, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018DC0246&from=EN;

"Review of Product Liability Rules: BEUC Position Paper", The European Consumer Organization: BEUC, accessed August 20, 2019, https://www.beuc.eu/publications/beuc-x-2017-039_csc_review_of_product_liability_rules.pdf;

"Robotics and artificial intelligence", House of Commons Science and Technology Committee, accessed June 14, 2019, https://publications.parliament.uk/pa/cm201617/cmselect/cmsctech/145/145.pdf;

"Special Eurobarometer 447 (regarding users concerned about data collected about them in the internet) European Commission, accessed July 5, 2019, http://ec.europa.eu/information_society/newsroom/image/document/2016-24/ebs_447_en_16136.pdf.

"State-of-the-Art Report | Algorithmic decision-making", algo:aware, accessed July 5, 2019, https://www.algoaware.eu/wp-content/uploads/2018/08/AlgoAware-State-of-the-Art-Report.pdf;

"The National Artificial Intelligence R&D Strategic Plan: 2019 update", Select Committee on Artificial Intelligence and National Science & Technology Council", accessed August 15, 2019, https://www.whitehouse.gov/wp-content/uploads/2019/06/National-AI-Research-and-Development-Strategic-Plan-2019-Update-June-2019.pdf;

"Understanding the challenges and opportunities of smart cities" Philips Lighting and SmartCitiesWorld, accessed June 11, 2019 http://www.lighting.philips.com/main/inspiration/smart-cities/smart-city-trends/smart-cities-world?origin=10_global_en_smartcities_pressrelease___scwnreport_70124000000WUc;

"White Paper on Artificial Intelligence at the service of citizens", Task Force on Artificial Intelligence of the Agency for Digital Italy, accessed June 11, 2019, https://ia.italia.it/assets/whitepaper.pdf;

"WJP Rule of Law Index 2019", World Justice Project, https://worldjusticeproject.org/sites/default/files/documents/WJP-ROLI-2019-Single%20Page%20View-Reduced.pdf;

"Work in the age of artificial intelligence: Four perspectives on the economy, employment, skills and ethics", Finish Ministry of Economic Affairs and Employment, accessed 10 June 2019, http://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/160980/TEMjul_21_2018_Work_in_the_age.pdf;

Directorate-General for Internal Policies of the Union and Directorate for Legislative Coordination and Conciliations Conciliations and Codecision Unit, *Handbook on the Ordinary Legislative Procedure* (European Parliament: Brussels, 2017), 11-25 and 37-39:

Directorate-General for Internal Policies of the Union and Directorate for Legislative Coordination and Conciliations Conciliations and Codecision Unit, *Activity Report on the*

Ordinary Legislative Procedure: 4 July 2014 - 31 December 2016 (8th parliamentary term) (European Parliament, Brussels: 2017), 19ff.;

European Commission, *Evaluation of Council Directive 85/374/EEC on the approximation of laws, regulations and administrative provisions of the Member States concerning liability for defective products* (Luxembourg: Publications Office of the European Union, 2018);

European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics, http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2017-0051+0+DOC+XML+V0//EN;

Ibero-American Data Protection Network, *General Recommendations for the Processing of Personal Data in Artificial Intelligence* (Juárez: Mexico, 2019);

Information Commissioner's Office, *Guide to the General Data Protection Regulation*, (London: ICO, May 2019)

US Department of Transportation, Preparing for the Future of Transportation: Automated Vehicle 3.0 (Washington, DC: US Department of Transportation, 2018;

World Intellectual Property Organization, *WIPO Technology Trends 2019: Artificial Intelligence* (Geneva: WIPO, 2019).

4. Case-Law

ECLI:EU:C:1968:6; Judgment of the ECJ of 16 June 1993, *France v. Commission*, Case C-325/91, ECLI:EU:C:1993:245;

Judgement of the ECJ of 5 April 1979, *Ratti*, Case C-148/78, ECLI:EU:C:1979:110

Judgment of EGC of 3 February 2017, *Minority SafePack - one million signatures for diversity in Europe*, Case T-646/13, ECLI:EU:T:2017:59;

Judgment of the Court of First Instance of 13 September 1995, *TWD Textilwerke Deggendorf GmbH v Commission,* Joined Cases T-244/93 and T-486/93, ECLI:EU:T:1995:160;

Judgment of the ECJ of 1 July 2008, *Turco v. Council*, Joined Cases C-39/05 P and C-52/05 P, ECLI:EU:C:2008:374;

Judgment of the ECJ of 1 October 2009, *Commission v. Council*, Case 370/07, ECLI:EU:C:2009:590;

Judgment of the ECJ of 1 October 2019, *Planet49*, Case C-673/17, ECLI:EU:C:2019:801;

Judgment of the ECJ of 10 July 2019, *Tietosuojavaltuutettu v. Jehovan todistajat — uskonnollinen yhdyskunta*, Case C–15/17, ECLI:EU:C:2018:551;

Judgment of the ECJ of 10 May 2001, *Henning Veedfald v Århus Amtskommune*, Case C-203/99, ECLI:EU:C:2001:258;

Judgment of the ECJ of 11 November 1997, *Eurotunnel SA v. SeaFrance*, Case C-408/95, ECLI:EU:C:1997:532, 35-39.;

Judgment of the ECJ of 12 September 2017, *Alexios Anagnostakis v. Commission,* C-589/15 P, ECLI:EU:C:2017:663;

Judgment of the ECJ of 13 July 1989, *Casa Fleischhandels v. Bundesanstalt für landwirtschaftliche Marktordnung*, Case 215/88, ECLI:EU:C:1989:331;

Judgment of the ECJ of 13 May 2014, *Google Spain v. Agencia Española de Protección de Datos*, Case C-131/12, ECLI:EU:C:2014:317;

Judgment of the ECJ of 14 April 2015, *Council v Commission*, Case C-409/13, ECLI:EU:C:2015:217;

Judgment of the ECJ of 14 February 2019, Sergejs Buivids v. Datu valsts inspekcija, Case C–345/17, ECLI:EU:C:2019:122;

Judgment of the ECJ of 15 May 1997, *TWD Textilwerke Deggendorf GmbH v Commission,* Case C-355/95 P, ECLI:EU:C:1997:241;

Judgment of the ECJ of 16 July 1992, *Parliament v. Council*, Case C-65/93 ECLI:EU:C:1992:325, 23-27;

Judgment of the ECJ of 16 June 1993, *France v. Commission*, C-325/91, ECLI:EU:C:1993:245;

Judgment of the ECJ of 17 October 2013, *Access Info Europe v. Council*, Case C-280/11 P, ECLI:EU:C:2013:671,

Judgment of the ECJ of 19 November 1998, *Nilsson*, Case 162/97, ECLI:EU:C:1998:554;

Judgment of the ECJ of 19 October 2016, *Patrick Breyer v. Bundesrepublik Deutschland*, Case C-582/14, ECLI:EU:C:2016:779;

Judgment of the ECJ of 21 December 2011, *Centre hospitalier universitaire de Besançon v. Thomas Dutrueux*; Case C-495/10, ECLI:EU:C:2011:869

Judgment of the ECJ of 21 December 2016, *Club Hotel Loutraki v. Commission*, C-131/15 P, ECLI:EU:C:2016:989;

Judgment of the ECJ of 24 November 2005, *Deutsches Milch-Kontor GmbH v Hauptzollamt Hamburg-Jonas*, Case C-136/04, ECLI:EU:C:2005:716.

Judgment of the ECJ of 24 September 2019, *Google v. CNIL*, Case C-507/17, ECLI:EU:C:2019:772;

Judgment of the ECJ of 25 November 1998, *Giuseppe Manfredi v. Regione Puglia*, Case C–308/97, ECLI:EU:C:1998:566;

Judgment of the ECJ of 26 June 2001, *The Queen v Secretary of State for Trade and Industry, ex parte Broadcasting, Entertainment, Cinematographic and Theatre Union*, Case C-173/99, ECLI:EU:C:2001:356;

Judgment of the ECJ of 27 November 2007, *C.*, C-435/06, ECLI:EU:C:2007:714;

Judgment of the ECJ of 29 July 2019, *Fashion Id.*, Case C-40/17, ECLI:EU:C:2019:629;

Judgment of the ECJ of 29 October 1980, *SA Roquette Frères v. Conselho*, Case 138/79, ECLI:EU:C:1980:249;

Judgment of the ECJ of 4 July 1963, *Germany v. Commission,* 24/15, ECLI:EU:C:1963:14;

Judgment of the ECJ of 4 September 2018, *ClientEarth v Commission*, Case C-57/16 P, ECLI:EU:C:2018:660;

Judgment of the ECJ of 5 June 2019, *Wirtschaftsakademie*, Case C-210/16, ECLI:EU:C:2018:388;

Judgment of the ECJ of 8 February 1968, *Fonderie Acciaierie Giovanni Mandelli v. Commission of the European Communities*, Case 3/67,

Judgment of the EGC of 10 May 2017, *Stop TTIP (Efler v. Commission)*, Case T-754/14, ECLI:EU:T:2017:323;

Judgment of the EGC of 13 November 2015, *ClientEarth v Commission*, Case T-424/14, ECLI:EU:T:2015:848;

Judgment of the EGC of 22 March 2011, *Access Info Europe v. Council*, Case T-233/09, ECLI:EU:T:2011:105;

Judgment of the EGC of 22 March 2018, *De Capitani v Parliament*, Case T-540/15, ECLI:EU:T:2018:167;

Judgment of the EGC of 24 September 2019, *Romania v. Commission*, Case T-391/17, ECLI:EU:T:2019:672.

Judgment of the EGC of 30 September 2015, *Alexios Anagnostakis v. Commission*, Case T-450/12, ECLI:EU:T:2015:739;

Judgment of the Supreme Court of the United States of 21 January 2010, *Citizens United v. the Federal Election Commission*;

Judgment of the Supreme Court of the United States of 22 June 1964, *Jacobellis v. Ohio*;

Judgment of the Supreme Court of the United States of 30 June 2014, *Burwell v. Hobby Lobby Stores, Inc*;

Opinion of the Advocate General delivered on 19 December 2019, *Fashion ID v. Verbraucherzentrale,* Case- C-40/17; ECLI:EU:C:2018:1039;

Opinion of the Advocate General delivered on 24 October 2017, *Wirtschaftsakademie*, Case C-210/16; ECLI:EU:C:2017:796;

Opinion of the Advocate General delivered on 27 May 1997, *Eurotunnel SA v. SeaFrance*, Case C-408/95, ECLI:EU:C:2014:2470;

5. Online Resources

"10 imperatives for Europe in the age of AI and automation", McKinsey Global Institute, accessed March 20, 2019, https://www.mckinsey.com/featured-insights/europe/ten-imperatives-for-europe-in-the-age-of-ai-and-automation;

"A GAMEBOY supercomputer", Kamil Rocki, accessed 15 September 2019, https://towardsdatascience.com/a-gameboy-supercomputer-33a6955a79a4;

"A Look at Germany's AI Strategy", Fabian Schmidt, accessed 15 July, 2019, https://iiot-world.com/artificial-intelligence/a-look-at-germanys-ai-strategy/;

"A major flaw in Google's algorithm allegedly tagged two black people's faces with the word 'gorillas'", Molly Mulshine, accessed June 10, 2019, https://www.businessinsider.com/google-tags-black-people-as-gorillas-2015-7;

"A Method for Detecting Bias in Search Rankings, with Evidence of Systematic Bias Related to the 2016 Presidential Election", Robert Epstein et al., accessed August 20, 2019, https://aibrt.org/downloads/EPSTEIN_et_al_2017-SUMMARY-A_Method_for_Detecting_Bias_in_Search_Rankings-EMBARGOED_until_March_14_2017.pdf.

"A new video shows a Tesla driver who appears to be asleep while driving down the Massachusetts Turnpike", Lisa Eadicicco, accessed September 11, 2019, https://www.businessinsider.com/tesla-driver-asleep-while-driving-down-massachusetts-turnpike-video-2019-9;

"A passenger jet pilot swerved to avoid drone near Gatwick Airport", Rob Picheta, accessed August 30, 2019, https://edition.cnn.com/2019/08/28/uk/gatwick-drone-near-miss-scli-gbr-intl/index.html;

"A Sleeping Tesla Driver Highlights Autopilot's Biggest Flaw", Alex Davies, accessed September 11, 2019, https://www.wired.com/story/tesla-sleeping-driver-dui-arrest-autopilot/;

"Actionable Auditing: Investigating the Impact of Publicly Naming Biased Performance Results of Commercial AI Products", Inioluwa Deborah Raji and Joy Buolamwini, accessed June 12, 2019, http://www.aies-conference.com/wp-content/uploads/2019/01/AIES-19_paper_223.pdf;

"AI and the Kratt momentum", e-Estonia, accessed June 12, 2019, https://e-estonia.com/ai-and-the-kratt-momentum/;

"AI Made in Germany — The German Strategy for Artificial Intelligence", C. Koch, accessed 15 July, 2019, https://towardsdatascience.com/ai-made-in-germany-the-german-strategy-for-artificial-intelligence-e86e552b39b6;

"AI Policy – China", Future of Life Institute, accessed June 29, 2019, https://futureoflife.org/ai-policy-china/;

"AI Policy – Japan", Future of Life Institute, accessed August 13, 2019, https://futureoflife.org/ai-policy-japan/;

"AI Robots Learning Racism, Sexism and Other Prejudices From Humans, Study Finds", Ian Johnston, accessed August 5, 2019, https://www.independent.co.uk/life-style/gadgets-and-tech/news/ai-robots-artificial-intelligence-racism-sexism-prejudice-bias-language-learn-from-humans-a7683161.html;

"Airports scramble to handle drone incidents", Matt McFarland, accessed August 30, 2019, https://edition.cnn.com/2019/03/05/tech/airports-drones/index.html;

"Algorithmic Justice League", accessed June 10, 2019, https://www.ajlunited.org/;

"Amazon Is Pushing Facial Technology That a Study Says Could Be Biased", Natasha Singer, accessed June 12, 2019, https://www.nytimes.com/2019/01/24/technology/amazon-facial-technology-study.html;

"Amazon offers cautionary tale of AI-assisted hiring", Andrew Hill, accessed June 11, 2019, https://www.ft.com/content/5039715c-14f9-11e9-a168-d45595ad076d;

"Amazon reportedly scraps internal AI recruiting tool that was biased against women", James Vincent, accessed June 11, 2019, https://www.theverge.com/2018/10/10/17958784/ai-recruiting-tool-bias-amazon-report;

"Amazon scrapped 'sexist AI' tool", BBC News, accessed June 11, 2019, https://www.bbc.com/news/technology-45809919;

"Amazon scraps secret AI recruiting tool that showed bias against women", Jeffrey Dastin, accessed June 11, 2019, https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G;

"Amazon shareholders reject ban on selling face recognition software to police", Laurence Dodds, accessed June 12, 2019, https://www.telegraph.co.uk/technology/2019/05/22/amazon-shareholders-reject-ban-selling-face-recognition-software/;

"Amazon shareholders reject facial recognition sale ban to governments", Zack Whittaker, accessed June 12, 2019, https://techcrunch.com/2019/05/22/amazon-reject-facial-recognition-proposals/;

"Amazon shareholders support selling face recognition tech to police", France24, accessed June 12, 2019, https://www.france24.com/en/20190522-amazon-shareholders-support-selling-face-recognition-tech-police;

"Amazon's Secret AI Hiring Tool Reportedly 'Penalized' Resumes With the Word 'Women's'", Rhett Jones, accessed June 11, 2019, https://gizmodo.com/amazons-secret-ai-hiring-tool-reportedly-penalized-resu-1829649346;

"An insider's view of Lisbon's rapidly growing tech scene", Clara Armand-Delille, accessed June 12, 2019, https://tech.eu/free/21951/an-insiders-view-of-lisbons-rapidly-growing-tech-scene/;

"Appointment of members of the High-Level Expert Group on the Impact of the Digital Transformation on EU Labour Markets", European Commission, accessed July 17, 2019, https://ec.europa.eu/digital-single-market/en/news/appointment-members-high-level-expert-group-impact-digital-transformation-eu-labour-markets;

"Artificial Intelligence for the American People: AI with American Values", United States Government, accessed August 15, 2019, https://www.whitehouse.gov/ai/ai-american-values/;

"Artificial Intelligence in Amsterdam, the City of Freedom", Simona Nickman, accessed June 11, 2019, https://medium.com/cityai/artificial-intelligence-in-amsterdam-the-city-of-freedom-83406e866e7e;

"Artificial Intelligence in Belgium and Luxembourg: How 277 major European companies benefit from AI", Microsoft, accessed June 10, 2019, https://info.microsoft.com/WE-DIGTRNS-CNTNT-FY19-09Sep-27-ArtificialIntelligenceinBelgium-MGC0003166_01Registration-ForminBody.html?wt.mc_id=AID732606_QSG_280351;

"Artificial Intelligence in Your Toilet. Yes, Really!", "Bernard Marr", accessed September 10, 2019, https://www.forbes.com/sites/bernardmarr/2019/05/20/artificially-intelligent-toilets-yes-they-are-here/#7f3c1585626d;

"Artificial Intelligence: What Is Reinforcement Learning - A Simple Explanation & Practical Examples", Bernard Marr, access January 10, 2019, https://www.forbes.com/sites/bernardmarr/2018/09/28/artificial-intelligence-what-is-reinforcement-learning-a-simple-explanation-practical-examples/#59db807f139c;

"Basic Questions", John McCarthy, accessed February, 5, 2019, http://www-formal.stanford.edu/jmc/whatisai/node1.html;

"Beijing AI Principles", Tsinghua University et al., , accessed June 29, 2019, https://www.baai.ac.cn/blog/beijing-ai-principles;

"Best Practices for Setting Up Meta Robots Tags and Robots.txt", Sergey Grybniak, accessed in June 4, 2019, https://www.searchenginejournal.com/best-practices-setting-meta-robots-tags-robots-txt/188655/#close;

"Bonjour smart alarm clock mysteriously killed off, despite $1m crowdfunding success", Alistair Charlton, accessed September 10, 2019, https://www.gearbrain.com/bonjour-smart-alarm-clock-killed-2626720300.html;

"Boris Johnson vows to ditch backstop and scale up no-deal plans", Peter Walker, accessed June 13, 2019, https://www.theguardian.com/politics/2019/jul/25/boris-johnson-vows-to-completely-ditch-brexit-backstop;

"Boris Johnson says he awaits EU move on Brexit delay after parliament defeat", Alasdair Sandford, accessed October 28, 2019, https://www.euronews.com/2019/10/23/watch-live-boris-johnson-addresses-uk-parliament-after-brexit-vote-defeat;

"Bulgaria's tax agency fined $3 million over data breach, will appeal", Reuters, accessed September 6, 2019, https://www.reuters.com/article/us-bulgaria-cybersecurity-fine/bulgarias-tax-agency-fined-3-million-over-data-breach-will-appeal-idUSKCN1VJ0YY.

"China Blocks Qualcomm's Attempt to Buy a Dutch Chipmaker", Klint Finley, accessed August 1, 2019, https://www.wired.com/story/china-blocks-qualcomms-attempt-to-buy-a-dutch-chipmaker/;

"China has started ranking citizens with a creepy 'social credit' system — here's what you can do wrong, and the embarrassing, demeaning ways they can punish you", Alexandra Ma, accessed July 15, 2019, https://www.businessinsider.com/china-social-credit-system-punishments-and-rewards-explained-2018-4;

"China seeks semiconductor security in wake of ZTE ban", Edward White, accessed June 17, 2019, https://www.ft.com/content/a1a5f0fa-63f7-11e8-90c2-9563a0613e56;

"China's edge in the tech race is vast amounts of data", Shafi Musaddique, accessed July 10, 2019, https://www.cnbc.com/2018/11/30/chinas-edge-in-the-tech-race-is-vast-amounts-of-data.html;

"China's Orwellian Social Credit Score Isn't Real", Jamie Horsley, accessed July 16, 2019, https://foreignpolicy.com/2018/11/16/chinas-orwellian-social-credit-score-isnt-real/;

"China's greatest natural resource may be its data", "Kai-Fu Lee", accessed July 16, 2019, https://www.cbsnews.com/news/60-minutes-ai-chinas-greatest-natural-resource-may-be-its-data-2019-07-14/;

"Chinese Interests Take a Big Seat at the AI Governance Table", Jeffrey Ding, Paul Triolo, and Samm Sacks, accessed June 10, 2019, https://www.newamerica.org/cybersecurity-initiative/digichina/blog/chinese-interests-take-big-seat-ai-governance-table/;

"Clinton Couldn't Win Over White Women", Clare Malone, accessed March, 20 2019, https://fivethirtyeight.com/features/clinton-couldnt-win-over-white-women;

"Cold water hits China's AI industry", Louise Lucas, accessed July 10, 2019, https://www.ft.com/content/973bfc08-a15f-11e9-a282-2df48f366f7d;

"Data trusts: why we are interested", Jack Hardinges, accessed June 7, 2019, https://theodi.org/article/what-is-a-data-trust/#1527168650599-ae3e3b8c-e22a62d2-2d92;

"Declaration of Cooperation on AI – Signed by Austria, Belgium, Bulgaria, Czech Republic, Denmark, Estonia, Finland, France, Germany, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Slovakia, Slovenia, Spain, Sweden, UK and Norway", EU Member States, accessed March 5, 2019, https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50951;

"Defense Department pledges billions toward artificial intelligence research", Drew Harwell, accessed August 12, 2019, https://www.washingtonpost.com/technology/2018/09/07/defense-department-pledges-billions-toward-artificial-intelligence-research/;

"Defining a 'data trust'"; Jack Hardinges and Peter Wells, accessed June 7, 2019, https://theodi.org/article/defining-a-data-trust/.

"Don't Get Your Valentine an Internet-Connected Sex Toy", Emily Dreyfuss, accessed July 10, 2019, "Autoblow AI is a sex toy that promises 'surprise'", Daniel Cooper, accessed

September 30, 2019, https://www.engadget.com/2019/09/27/autoblow-ai-deep-learning-sex-toy/;

"Donald Trump hits out at Facebook's Libra and bitcoin" Hannah Murphy, accessed August 12, 2019, https://www.ft.com/content/57692326-a452-11e9-974c-ad1c6ab5efd1;

"Donald Trump: Apple should make products in the US to avoid tariffs", The Guardian Staff, accessed August 12, 2019, https://www.theguardian.com/us-news/2018/sep/09/donald-trump-apple-should-make-products-in-the-us-to-avoid-tariffs;

"Edición del lunes, 09 marzo 1998", La Vanguardia, accessed in June 4, http://hemeroteca.lavanguardia.com/preview/2013/02/27/pagina-13/33837533/pdf.html;

"EU agrees to January 31 Brexit extension", Jacopo Barigazzi and James Randerson, accessed October 28, 2019, https://www.politico.eu/article/eu-agrees-to-january-31-brexit-extension/;

"EU Court Condemns the EU Legislative Process for Lack of Transparency: Time to Open Up?", Massimo Frigo, accessed December 10, 2018, http://opiniojuris.org/2018/03/27/33507/;

"EU Member States sign up to cooperate on Artificial Intelligence", European Institutions, accessed March 5, 2019, https://ec.europa.eu/digital-single-market/en/news/eu-member-states-sign-cooperate-artificial-intelligence;

"Europe's new data protection rules export privacy standards worldwide", Mark Scott and Laurens Cerulus, accessed April 10, 2019, https://www.politico.eu/article/europe-data-protection-privacy-standards-gdpr-general-protection-data-regulation/;

"European Accessibility Act: final steps on the European level – first steps on the national level", European Union of the Deaf, accessed June 20, 2019, https://www.eud.eu/news/european-accessibility-act-final-steps-european-level-first-steps-national-level/;

"European Accessibility Act", European Commission, accessed June 20, 2019, https://ec.europa.eu/social/main.jsp?catId=1202 ;

"European Commission adopts adequacy decision on Japan, creating the world's largest area of safe data flows", European Commission, accessed February 12, 2019, https://europa.eu/rapid/press-release_IP-19-421_en.htm;

"European officials draft radical plan to take on Trump and U.S. tech companies", Bjarke Smith-Meyer et al., accessed August 22, 2019, https://www.politico.com/story/2019/08/22/europe-plan-trump-tech-companies-1472326.

"Exclusive: Brussels eyes €100B wealth fund for 'European champions'", Bjarke Smith-Meyer, Lili Bayer And Jakob Hanke, accessed August 22, 2019, https://www.politico.eu/article/exclusive-european-commission-leaked-plans/;

"Expert group to the EU Observatory on the Online Platform Economy", accessed in October 13, 2019, https://ec.europa.eu/digital-single-market/en/expert-group-eu-observatory-online-platform-economy;

"Explainable Artificial Intelligence (XAI)", DARPA, accessed August 15, 2019, https://www.darpa.mil/program/explainable-artificial-intelligence;

"Facebook is Finding Problems With Artificial Intelligence Too", Tom Simonite, accessed June 10, 2019, https://www.wired.com/story/facebook-finding-problems-artificial-intelligence-too/.

"Fact check: Trump falsely claims Google 'manipulated' millions of 2016 votes", Daniel Dale, accessed August 20, 2019, https://edition.cnn.com/2019/08/19/politics/trump-google-manipulated-votes-claim/index.html;

"Finland's grand AI experiment: Inside Finland's plan to train its population in artificial intelligence", Janosch Delcker, accessed 10 June 2019, https://www.politico.eu/article/finland-one-percent-ai-artificial-intelligence-courses-learning-training/;

"Flickr faces complaints over 'offensive' auto-tagging for photos", Alex Hern, accessed June 10, 2019, https://www.theguardian.com/technology/2015/may/20/flickr-complaints-offensive-auto-tagging-photos;

"Flickr Fixing 'Racist' Auto-Tagging Feature After Black Man Mislabeled 'Ape'", Michael Zhang, accessed June 10, 2019, https://petapixel.com/2015/05/20/flickr-fixing-racist-auto-tagging-feature-after-black-man-mislabeled-ape/;

"Flickr's new auto-tags are racist and offensive", David Goldman, accessed June 10, https://money.cnn.com/2015/05/21/technology/flickr-racist-tags/;

"France wants to become an artificial intelligence hub: attracting talent, fostering public research and AI startups", Romain Dillet, accessed May 7, 2019, https://techcrunch.com/2018/03/29/france-wants-to-become-an-artificial-intelligence-hub/;

"From 'Crooked Hillary' to 'Little Marco,' Donald Trump's Many Nicknames", Paola Chavez and Veronica Stracqualursi, accessed March 10, 2019, http://abcnews.go.com/Politics/crooked-hillary-marco-donald-trumps-nicknames/story?id=39035114;

"Fully Sleeping" Tesla Driver Cruises 30 Miles on Autopilot", Kristin Houser, accessed September 11, 2019, https://futurism.com/the-byte/sleeping-tesla-driver-autopilot;

"Gatwick Airport: Drones ground flights", BBC News, accessed August 30, 2019, https://www.bbc.com/news/uk-england-sussex-46623754;

"Gatwick returns to normality but drone threat remains", Jamie Grierson, accessed August 30, 2019, https://www.theguardian.com/world/2019/jan/04/gatwick-returns-to-normality-but-drone-threat-remains;

"GDP by Member State, Eurostat", accessed July 5, 2019, http://appsso.eurostat.ec.europa.eu/nui/submitViewTableAction.do.;

"Google 'fixed' its racist algorithm by removing gorillas from its image-labeling tech", James Vincent, accessed June 10, 2019, https://www.theverge.com/2018/1/12/16882408/google-racist-gorillas-photo-recognition-algorithm-ai;

"Google Acquires Artificial Intelligence Startup DeepMind For More Than $500M", Catherine Shu, access March 20, 2019, https://techcrunch.com/2014/01/26/google-deepmind/;

"Google apologises for Photos app's racist blunder", BBC News, accessed June 10, 2019, https://www.bbc.com/news/technology-33347866;

"Google says sorry for racist auto-tag in photo app", Jana Kasperkevic, accessed June 10, 2019, https://www.theguardian.com/technology/2015/jul/01/google-sorry-racist-auto-tag-photo-app;

"Guide to EU decision-making and justice and home affairs after the Treaty of Lisbon", Steve Peers, accessed December 13, 2019, http://www.statewatch.org/analyses/no-115-lisbon-treaty-decision-making.pdf;

"Here Are the Microsoft Twitter Bot's Craziest Racist Rants", Sophie Kleeman, accessed August 4, 2019, https://gizmodo.com/here-are-the-microsoft-twitter-bot-s-craziest-racist-ra-1766820160;

"Homo sapiens 100,000 years older than thought", Clive Cookson, accessed April 12, 2019, , https://www.ft.com/content/00a266c2-4ad3-11e7-919a-1e14ce4af89b;

"How a tiny country bordering Russia became one of the most tech-savvy societies in the world", Elizabeth Schulze, accessed June 12, 2019, https://www.cnbc.com/2019/02/08/how-estonia-became-a-digital-society.html;

"How France became the place to be for AI startups", TechStartups Team, accessed May 7, 2019, https://techstartups.com/2019/01/16/france-became-place-ai-startups/;

"How Jeb Bush Spent $130 Million Running for President With Nothing to Show for It", Nicholas Confessore and Sarah Cohen, accessed March 10, 2019, https://www.nytimes.com/2016/02/23/us/politics/jeb-bush-campaign.html?_r=1;

"How Microsoft is Using AI to Tackle Fake News", James O Malley, accessed August 2, 2019, http://www.gizmodo.co.uk/2018/05/how-microsoft-is-using-ai-to-tackle-fake-news/;

"Huawei Can't Guarantee Mate 30 Will Ship With Android, But Will Be 'Ready' With Alternative OS", Ben Sin, accessed July 30, 2019, https://www.forbes.com/sites/bensin/2019/07/30/huawei-cant-guarantee-mate-30-will-ship-with-android-but-will-be-ready-with-alternative-os/#354066445d09;

"Intention to fine British Airways £183.39m under GDPR for data breach", ICO, accessed September 6, 2019, https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/07/ico-announces-intention-to-fine-british-airways/;

"International Digital Economy and Society Index 2018, Tech4i2 (Professor Paul Foley, Dr David Sutton, Ian Wiseman, Lawrence Green and Jake Moore) and European Commission, accessed August 10, 2019, http://ec.europa.eu/newsroom/dae/document.cfm?doc_id=50224;

"Irish minister says Boris Johnson's Brexit stance 'quite alarming'", Lisa O'Carrol, accessed June 13, 2019, https://www.theguardian.com/politics/2019/jul/25/irish-minister-says-boris-johnsons-brexit-stance-quite-alarming;

"Japan falling behind in artificial intelligence, warns SoftBank founder", Kana Inagaki, accessed August 13, 2019, https://www.ft.com/content/cab0936c-a940-11e9-984c-fac8325aaa04; Toyoaki Nishida,

"Meet the World's Most Valuable AI Startup: China's SenseTime", Bernard Marr, accessed July 5, 2019, https://www.forbes.com/sites/bernardmarr/2019/06/17/meet-the-worlds-most-valuable-ai-startup-chinas-sensetime/#7bc932e7309f;

"Microsoft improves facial recognition technology to perform well across all skin tones, genders", John Roach, accessed June 12, 2019, https://blogs.microsoft.com/ai/gender-skin-tone-facial-recognition-improvement/;

"Microsoft's neo-Nazi sexbot was a great lesson for makers of AI assistants", Rachel Metz accessed August 4, 2019, https://www.technologyreview.com/s/610634/microsofts-neo-nazi-sexbot-was-a-great-lesson-for-makers-of-ai-assistants/;

"Microsoft's politically correct chatbot is even worse than its racist one", accessed August 5, 2019, https://qz.com/1340990/microsofts-politically-correct-chat-bot-is-even-worse-than-its-racist-one/;

"Microsoft's racist chatbot returns with drug-smoking Twitter meltdown"; Samuel Gibbs, accessed August 5, 2019, https://www.theguardian.com/technology/2016/mar/30/microsoft-racist-sexist-chatbot-twitter-drugs;

"Mitigating Bias in AI Models", Ruchir Puri, accessed June 12, 2019, https://www.ibm.com/blogs/research/2018/02/mitigating-bias-ai-models/;

"Move over AlphaGo: AlphaZero taught itself to play three different games", Jennifer Ouellette, access March 15, 2019, https://arstechnica.com/science/2018/12/move-over-alphago-alphazero-taught-itself-to-play-three-different-games/.

"New rules on contracts for the supply of digital content and digital services", Cătălin Grigorescu and Diana Gavril, accessed June 5, 2019, https://www.bpv-grigorescu.com/publications/legal-tax-alerts/new-rules-on-contracts-for-the-supply-of-digital-content-and-digital-services/;

"Next European Commission takes aim at AI", Laura Kayali, accessed September 10, 2019, https://www.politico.eu/article/ai-data-regulator-rules-next-european-commission-takes-aim/;

"Not content with digital – new protections for EU consumers of digital content and services", Edward Turtle and Evangelia Nitti, accessed June 5, 2019, https://products.cooley.com/2019/03/15/not-content-with-digital-new-protections-for-eu-consumers-of-digital-content-and-services/;

"NSF Program on Fairness in Artificial Intelligence (AI) in Collaboration with Amazon (FAI)", National Science Foundation, accessed August 15, 2019, https://www.nsf.gov/funding/pgm_summ.jsp?pims_id=505651.

"NVIDIA's 'kitchen manipulator' is the ultimate robot chef", designboom, accessed September 10, 2019, https://www.designboom.com/technology/nvidia-kitchen-manipulator-robot-chef-cobot-15-01-2019/;

"On Recent Research Auditing Commercial Facial Analysis Technology", Ali Alkhatib et al., accessed June 12, 2019, https://medium.com/@bu64dcjrytwitb8/on-recent-research-auditing-commercial-facial-analysis-technology-19148bda1832;

"Open Letter on Privacy", VV.AA., accessed September 10, 2019, https://s3.amazonaws.com/brt.org/BRT-CEOLetteronPrivacy-2.pdf.;

"Party Brands, Elections, and Presidential-Congressional Relations", David R. Jones, accessed December 11, 2018, https://www.baruch.cuny.edu/wsas/academics/political_science/documents/PartyBrandsElectionsandPresidentialCongressionalRelations.pdf;

"Poll: Trump administration edges media in voter trust", Cristiano Lima, accessed March 10, 2019, http://www.politico.com/story/2017/02/trump-media-trust-poll-fox-news-235168;

"Response: Racial and Gender bias in Amazon Rekognition — Commercial AI System for Analyzing Faces.", Joy Buolamwini, accessed June 12, 2019, https://medium.com/@Joy.Buolamwini/response-racial-and-gender-bias-in-amazon-rekognition-commercial-ai-system-for-analyzing-faces-a289222eeced;

"Security Breach Notification Laws", National Conference of State Legislatures, accessed October 10, 2019, http://www.ncsl.org/research/telecommunications-and-information-technology/security-breach-notification-laws.aspx;

"Shutdown of the Federal Government: Causes, Effects, and Process", Kevin R. Kosar, accessed December 11, 2018, http://assets.thefiscaltimes.com/TFT2_20101228/App_Data/MediaFiles/1/B/1/%7B1B124168-264B-4686-8E6F-DE0D5F0E097E%7DShutdown%20background.pdf;

"Shutdown of the Federal Government: Causes, Processes, and Effects", Clinton T. Brass, accessed December 10, 2018, http://digitalcommons.ilr.cornell.edu/cgi/viewcontent.cgi?article=2182&context=key_workplace; "Shutdown of the Federal Government: Causes, Effects, and Process",

"Silicon Valley finally pushes for data privacy laws at Senate hearing", Dan Tynan, accessed April 5, 2019, https://www.theguardian.com/technology/2018/sep/26/silicon-valley-senate-commerce-committee-data-privacy-regulation;

"Some Thoughts on Facial Recognition Legislation", Michael Punke, accessed June 12, https://aws.amazon.com/blogs/machine-learning/some-thoughts-on-facial-recognition-legislation/?;

"Statement: Intention to fine Marriott International, Inc more than £99 million under GDPR for data breach", ICO, accessed September 6, 2019, https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2019/07/statement-intention-to-fine-marriott-international-inc-more-than-99-million-under-gdpr-for-data-breach/;

"Successful Stories", e-Estonia, accessed June 12, 2019, https://e-estonia.com/; https://www.cnbc.com/2019/02/08/how-estonia-became-a-digital-society.html;

"Sweden set to become global leader in artificial intelligence", The Swedish Trade and Investment Council, accessed June 10, 2019, https://www.business-sweden.se/en/Invest/industries/Data-Centers-By-Sweden/news-and-downloads/investment-news/sweden-set-to-become-global-leader-in-artificial-intelligence/;

"Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter", Ellie Hunt, accessed August 5, 2019, https://www.theguardian.com/technology/2016/mar/24/tay-microsofts-ai-chatbot-gets-a-crash-course-in-racism-from-twitter;

"Tesla drivers are getting caught sleeping on Autopilot – blame people, not Autopilot", Fred Lambert, accessed September 11, 2019, https://electrek.co/2019/06/16/tesla-driver-caught-sleepingn-autopilot-blame/;

"The Accountability of AI — Case Study: Microsoft's Tay Experiment", Yuxi Liu, accessed August 4, 2019, https://chatbotslife.com/the-accountability-of-ai-case-study-microsofts-tay-experiment-ad577015181f;

"The Best of AI in Japan — Prologue", AI Magazine, 33(2): 108-111; "Japan aims to produce 250,000 AI experts a year", Minako Yamashita, accessed August 12, 2019, https://asia.nikkei.com/Economy/Japan-aims-to-produce-250-000-AI-experts-a-year;

"The Chinese AI innovation chasm", Michael R. Wade and Amanda Bris, accessed July 12, 2019, https://www.imd.org/research-knowledge/articles/the-chinese-AI-innovation-chasm/;

"The CNIL's restricted committee imposes a financial penalty of 50 Million euros against GOOGLE LLC, CNIL, accessed September 6, 2019, https://www.cnil.fr/en/cnils-restricted-committee-imposes-financial-penalty-50-million-euros-against-google-llc;

"The conclusion of contracts by software agents in the eyes of the law", Tina Balke and Torsten Eymann, accessed April 6, 2019, https://www.researchgate.net/publication/221456172_The_conclusion_of_contracts_by_software_agents_in_the_eyes_of_the_law/link/0fcfd50adf322c7aa1000000/download;

"The European High-Performance Computing Joint Undertaking – EuroHPC", European Commission, accessed July 17, 2019, https://ec.europa.eu/digital-single-market/en/eurohpc-joint-undertaking;

"The Impact of the EU's New Data Protection Regulation on AI", Nick Wallace and Daniel Castro, accessed August 9, 2019, http://www2.datainnovation.org/2018-impact-gdpr-ai.pdf;

"The legacy of the Reverend Bayes", Mathematical Association of America, access March 18, 2019, https://www.maa.org/external_archive/devlin/devlin_2_00.html;

"The Next Big Privacy Hurdle: Teaching AI to Forget", Darren Shou, accessed September 7, 2019, https://www.wired.com/story/the-next-big-privacy-hurdle-teaching-ai-to-forget/;

"The Slow Birth of Agriculture", Heather Pringle, accessed April 12, 2019, https://science.sciencemag.org/content/282/5393/1446;

"The Special Legislative Procedures: Consent", University of Portsmouth European Studies Hub, accessed December 13, 2019, http://hum.port.ac.uk/europeanstudieshub/learning/module-2-understanding-eu-policy-making/the-special-legislative-procedures/;

"The State of European Tech 2018", Atomico, accessed June 13, 2019, https://2018.stateofeuropeantech.com/;

"The story of AlphaGo so far", DeepMind, access March 15, 2019, https://deepmind.com/research/alphago/;

"This company specialises in talking, AI-powered sex dolls", BBC News, accessed July 10, 2019, https://www.bbc.com/reel/video/p06f6xn2/this-company-specialises-in-talking-ai-powered-sex-dolls;

"This Guy Made a 'Game Boy Supercomputer' That Can Handle 1 Billion Frames Per Second", Daniel Oberhaus, accessed 15 September 2019, https://www.vice.com/en_us/article/qvqamb/this-guy-made-a-game-boy-supercomputer-that-can-handle-1-billion-frames-per-second;

"Thoughts on Recent Research Paper and Associated Article on Amazon Rekognition", Matt Wood, accessed June 12, 2019, https://aws.amazon.com/blogs/machine-learning/thoughts-on-recent-research-paper-and-associated-article-on-amazon-rekognition/?;

"Top 15 Deep Learning applications that will rule the world in 2018 and beyond", Vartul Mittal, accessed June 10, 2019, https://medium.com/breathe-publication/top-15-deep-learning-applications-that-will-rule-the-world-in-2018-and-beyond-7c6130c43b01;

"Trump Attacks Amazon, Saying It Does Not Pay Enough Taxes", Michael D. Shear et al., accessed August 12, 2019, https://www.nytimes.com/2018/03/29/us/politics/trump-amazon-taxes.html;

"Trump calls for Apple boycott", Jeremy Diamond, accessed August 12, 2019, https://edition.cnn.com/2016/02/19/politics/donald-trump-apple-boycott/index.html;

"Trump claims Google is suppressing positive news about him and 'will be addressed", James Vincent, August 12, 2019, https://www.theverge.com/2018/8/28/17790164/president-trump-google-left-wing-bias-claims;

"Trump criticizes Twitter in tweet, urges 'fairer' social media", Makini Brice and Susan Heavey, accessed August 12, 2019, https://www.reuters.com/article/us-usa-trump-twitter/trump-criticizes-twitter-in-tweet-urges-fairer-social-media-idUSKCN1RZ171;

"Trump had a 'very good meeting' with Apple's Tim Cook about tariffs on iPhones, devices", Dalvin Brown, accessed August 20, 2019, https://eu.usatoday.com/story/tech/2019/08/19/trump-says-apple-ceo-tim-cook-makes-compelling-case-against-tariffs/2049030001/;

"Trump orders U-turn over sanctions against Chinese telecoms group", Sam Fleming and Shawn Donnan, accessed June 17, 2019, https://www.ft.com/content/8fc1b404-56c4-11e8-b8b2-d6ceb45fa9d0;

"Trump says he'll ease Huawei restrictions, but no one's sure how", Colin Lecher, accessed July 30, 2019, https://www.theverge.com/2019/7/3/20679998/trump-huawei-trade-announcement-restrictions;

"Trump says he's 'watching Google very closely' after meeting with CEO", Colin Lecher, accessed August 12, 2019, https://www.theverge.com/2019/8/6/20756734/trump-google-anti-conservative-bias-claims-tweets;

"Trump tweets support for far-right figures banned by Facebook", Brian Stelter and Oliver Darcy, accessed August 12, 2019, https://edition.cnn.com/2019/05/04/tech/trump-social-media-twitter-facebook/index.html;

"Trump: Facebook, Twitter, Google are 'treading on very, very troubled territory and they have to be careful", Ryan Browne, accessed August 12, 2019, https://www.cnbc.com/2018/08/28/trump-accuses-google-of-rigging-search-results-in-favor-of-bad-coverage.html;

"U.K. Presses Ahead With Tech Tax Plans Despite Rising Tensions With Washington", Sam Shead, accessed August 10, 2019,

https://www.forbes.com/sites/samshead/2019/07/12/uk-presses-ahead-with-tech-tax-plans-despite-rising-tensions-with-washington/#100ad6116904.

"U.S. Card Fraud Losses Could Exceed $12B By 2020", Roger Aitken, access January, 23, 2019, https://www.forbes.com/sites/rogeraitken/2016/10/26/us-card-fraud-losses-could-exceed-12bn-by-2020/#276281c6d243;

"US companies banned from selling components to ZTE", Brian Heater, accessed June 17, 2019, https://techcrunch.com/2018/04/16/u-s-companies-banned-from-selling-components-to-zte/;

"US launches inquiry into French plan to tax tech giants", BBC News, accessed August 10, 2019, https://www.bbc.com/news/world-europe-48945828;

"US to probe proposed French tech tax, concerned it 'unfairly targets American companies'", CNBC, accessed August 10, 2019, https://www.cnbc.com/2019/07/11/us-to-probe-proposed-french-tech-tax.html; "UK braves US ire by pressing ahead with tax on tech companies", Chris Giles and Claer Barret, accessed August 10, 2019, https://www.ft.com/content/41069548-a3d8-11e9-974c-ad1c6ab5efd1;

"Video appears to show Tesla driver asleep at the wheel", The Guardian Staff, accessed September 11, 2019, https://www.theguardian.com/technology/2019/sep/10/video-appears-to-show-tesla-driver-asleep-at-the-wheel-car;

"Wallenberg AI, Autonomous Systems and Software Program (WASP)", Linköping University, accessed June 10, 2019, https://liu.se/en/research/wallenberg-artificial-intelligence-autonomous-systems-and-software-program;

"We really need to take accountability," Microsoft CEO on the 'Tay' chatbot", Charlie Moloney, accessed August 2, 2019, https://chatbotsmagazine.com/we-really-need-to-take-accountability-microsoft-ceo-on-the-tay-chatbot-2383ee83a6ba.

"Welcome to E-stonia, the world's most digitally advanced society", Matt Reynolds, accessed June 12, 2019, https://www.wired.co.uk/article/digital-estonia;

"What is Semi-Supervised Learning?", Nikki Castle, access February 20, 2019, https://www.datascience.com/blog/what-is-semi-supervised-learning;

"What People Get Wrong About China and Artificial Intelligence", Jonathan Vanian, accessed July 12, 2019, https://fortune.com/2019/07/09/china-data-artificial-intelligence/;

"What you may not understand about China's AI scene", Karen Hao, accessed June 13, 2019, https://www.technologyreview.com/f/613296/what-you-may-not-understand-about-chinas-ai-scene/;

"When it Comes to Gorillas, Google Photos Remains Blind", Tom Simonite, accessed June 10, 2019, https://www.wired.com/story/when-it-comes-to-gorillas-google-photos-remains-blind/;

"Why does Beijing suddenly care about AI ethics?", Will Knight, June 13, 2019, https://www.technologyreview.com/s/613610/why-does-china-suddenly-care-about-ai-ethics-and-privacy/;

"Women in Digital", European Commission, accessed August 10, 2019, https://ec.europa.eu/digital-single-market/en/women-ict;

"ZTE fined $1 billion", Danny Crichton, accessed June 17, 2019, https://techcrunch.com/2018/06/07/zte-fined-1 billion/?guccounter=1&guce_referrer_us=aHR0cHM6Ly93d3cuZ29vZ2xlLnB0Lw&guce_referrer_cs=kFeknkKSA2-T3R3ISgwUDA;

"What does Huawei's trade ban mean for your Huawei or Honor phone?", Dan Grabham, accessed October 25, 2019, https://www.pocket-lint.com/phones/news/huawei/148102-what-does-huawei-s-google-ban-mean-for-your-huawei-or-honor-phone;

"Huawei: Record 200 Million Devices Shipped As Android Future Resolved", Zak Doffman, accessed October 25, 2019, https://www.forbes.com/sites/zakdoffman/2019/10/24/huawei-record-200-million-devices-shipped-as-android-future-resolved/#34d2a2446f40.

Pedro Domingos; "Thomas Bayes", Encyclopedia Britannica, accessed March 17, 2019, https://www.britannica.com/biography/Thomas-Bayes.