SciencesPo
SCHOOL OF PUBLIC AFFAIRS

PUBLIC POLICY MASTER THESIS

May 2021

Algorithmic Transparency in the EU

Can Şimşek

Master's Thesis supervised by Florence G'sell

Second member of the Jury: Dominique Cardon

Master in European Affairs
Digital, New Technology and Public Policy

Abstract

Algorithmic systems that are utilized in decision making processes or taking autonomous decisions in various ways from curating the online content to enabling "artificial intelligence" appliances are proliferating. As the number of intellectual tasks that are delegated to such systems increases, policy makers raise concerns about accountability and fairness as well as the lack of transparency regarding such systems. To a certain extent, this lack of transparency is due to the complexity of techniques that create the "black box" algorithms. Yet, the main reason is the legal frameworks that veil the algorithmic systems. Namely, intellectual property rights, and "trade secrets" in particular, are balanced against the "public interest" in the EU just as in many other jurisdictions. Thereby, the hidden policy decisions in the "codes" cannot be scrutinized and the negative societal impacts of the algorithmic systems cannot be addressed by the regulatory authorities. Given that this situation is posing serious risks like undermining human rights, rule of law and democratic processes, there is a pressing need for further transparency and oversight. In this context, the EU introduced some regulations, starting with the GDPR, and plans to recalibrate algorithmic transparency within its new proposals such as the DSA and the AI Act that contain new modalities of transparency. Therefore, this thesis examines the European approach to algorithmic transparency and recommends amendments to the mentioned regulations in order to strengthen their enforcement structures and the wordings of some articles as well as adoption of stricter ex-ante rules and targeted transparency policies.

Key words

Algorithms, Transparency, European Union, Artificial Intelligence Act, Digital Services Act, Regulation

Table of Contents

CHAPTER 1: INTRODUCTION

1.0 Why should you read this research?

This research is a “must read” for the readers who are interested in AI regulation, algorithmic transparency, and legal philosophy concerning these topics since it aims to humbly contribute to the legislative efforts of the EU in a rather historical moment. In fact, it is -probably- the first thesis that covers the proposed “AI Act” of the European Union under its scope which will be the very first “AI Regulation” in the history if adopted. Moreover, it also contains a brief qualitative assessment of the proposed Digital Services Act which aims to change the structure of the digital realm by regulating online intermediary platforms.

Since “transparency” is the main regulatory tool that the EU prefers while dealing with the digital technologies, this research revolves around the modalities of transparency that are already in use and that are proposed within the initial drafts of the prospective regulations. For exploring the potential paths for the policy makers, the research begins by focusing on the evolution of the term “transparency” in the context of governance and tries to relate it to the evolution of the algorithmic systems in the historical context of the 20th century. Moreover, the interaction between the terms “transparency” and “algorithmic systems” is further investigated by evaluating the technical challenges that the “black box” algorithms pose concerning interpretability as well as the legal reasons of opacity. In this respect, the research focuses on the “trade secrets” as the prominent legal obstacle for the algorithmic transparency. Therefore, it might also be of interest for the readers who are interested in the intellectual property law and trade secrets in particular.

As an outcome of this research, the author provides an analysis of different forms of transparency depending on the receiver of the information (as individuals, public, trusted third parties or public authorities), the amount of the information (full transparency vs. limited transparency as intelligible explanation of specific decision or general decision models) and the timing (ex-ante or ex-post). From this broader perspective, the author suggests amendments to the GDPR, as well as the proposed drafts of the Digital Services Act and the AI Act. In this respect, creating EU Level enforcement bodies, clarifying the wordings of certain articles, introducing a right to reasonable inferences as well as introducing stricter ex-ante rules and mandatory pre-market authorization requirement for certain algorithmic systems are recommended. In addition to these recommendations, the author calls for the harmonization of “freedom of information laws” and the creation of repositories for publicizing the source codes and other relevant data of the algorithmic systems that are deployed by the public agencies and the private actors that are active in certain areas where such a disclosure could serve the “public interest.”

1.1 Introduction to the Inquiry

Transparency is a multi-dimensional concept with different connotations. Etymologically, the word has a physical origin since it is a combination of the Latin words *trans* (across, beyond; through) and *parere* (come in sight, appear; submit, obey) that appears as *transparere* (show light through) in early 15th century Medieval Latin. (Online Ethymology Dictionary, 2021) According to the Cambridge English Dictionary, transparency is "the characteristics of being easy to see through" (Cambridge Dictionary, 2021). In the context of governance and Law, the word "transparent" is often referred as a positive attribution signifying the accessibility of information. Moreover, "transparency" is increasingly seen as an essential component of a democratic regime. Arguably, this positive meaning had emerged in relation to the historical evolution of western democracies under the influence of Enlightenment ideals and reached today by taking a neoliberal turn during the late 20th century as a transparency pressure over the nation states. (Mehrpouya & Djelic, 2014). According to Hood, transparency concerning governance appears as an empirical ideal deriving from the fundamental epistemology of "eighteenth-century ideas about social science, that the social world should be made knowable by methods analogous to those used in the natural sciences" (2006, p. 8) Indeed, from the revolutions in France and America, to the passing of "freedom of information" legislations in Scandinavia; 18th century had been a milestone for the idea of holding governments accountable via transparency. In 1766, the Swedish Freedom of the Press Act (Tryckfrihetsförordningen) provided publishers the statutory right of accessing to government records and became the ancestor of modern "freedom of information laws". Adoption of freedom of information laws under which individuals can exercise their "right to know" through information requests was a huge step for transparency ideal although scopes of such laws have been often limited with the information that States held and excluded the private sector despite its increasing role in public services (Siraj, 2010).

Along with the transparency trend concerning "public affairs", a demand for opacity for the "personal" affairs materialized in Law during the 19th century. In a way, the increasing circulation of magazines and the invention of photography made lawyers and policy makers to discuss the other side of the medallion. During a debate about the 1819 press law in France, Royer-Collard used the expression "vie privée murée", referring to the private life conception that appeared in the 1791 Constitution which draws the limit of press freedom with "calumnies and insults relative to private life". (Wagner, 1971) Consequently, "private life" gained statutory recognition in France with the 1819 press law as it had to be "murée" (walled off). In 1867, the Paris Court of Appeals decided that the famous writer Alexander Dumas could withdraw his consent for giving publicity to his improper photographs. (Dumas v. Liebert, 1868) This case law was a milestone in terms of right to privacy since the Court held that individuals do have control over their personal information, in this case their personal photographs. Still, it was not until the end of the 19th century that "the right to privacy" was cast in cement. In establishing a full-fledged right to privacy doctrine, US Supreme Court Justice Louis Brandeis had been the pioneering figure. In fact, "the right to privacy" was coined after an article written in 1890 by him and Samuel D. Warren. In this article, they stated that ' 'the common law secures to each individual the right of determining, ordinarily, to what extent his thoughts, sentiments, and emotions shall be communicated to others." (Warren & Brandeis, 1890) Yet, twenty-three years after formulating the legal doctrine of right to privacy which provides individuals an opaque space where public cannot see through, Brandeis wrote: "Publicity is justly commended as a remedy for social and industrial diseases. Sunlight is said to be the best of disinfectants; electric light the most efficient policeman." (Brandeis, 1914). These words were written with a view to putting an end to the corrupted practices within the unstable US financial markets. Since the private sector gained more importance in national economies, economic affairs were no longer seen as private. Thus, differing from the transparency wave of the 18th century, legal doctrines revolving around transparency and political discourses of the 20th century were also targeting the private sector besides the States.

After two World Wars, international and supra-national organizations gained the utmost importance for upholding peace and safeguarding human rights. With the adoption of the Universal Declaration of Human Rights[1] in 1948, "information" finally became a well-established human right as being enshrined in the Article 19. Two years later, Council of Europe drafted the European Convention on Human Rights[2] (ECHR). The Article 10 (1) of the ECHR echoed the right to information within the right to freedom of expression as follows:

"Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers…"

In 1966, signature of the International Covenant on Civil and Political Rights[3] (ICCPR) further reaffirmed this right although it only came into force ten years after its signature. Article 19 (2) of the ICCPR writes that:

"Everyone shall have the right to freedom of expression; this right shall include freedom to seek, receive and impart information and ideas of all kinds, regardless of frontiers, either orally, in writing or in print, in the form of art, or through any other media of his choice."

Thereby, seeking, receiving, and imparting information was acknowledged as a part of the freedom of expression. Yet, it is noteworthy that the "right to information" is not an absolute right. Besides the possibility to derogate, at the very least, all these international human rights documents also incorporate a right to privacy and a right to private property which could limit this right to a certain extent. Even so, the mentioned documents rendered transparency as a rule that enables human rights and thus made opacity an exception.

[1] Universal Declaration of Human Rights. Adopted and proclaimed by General Assembly resolution 217 A (III) of 10 December 1948.

[2] Convention for the Protection of Human Rights and Fundamental Freedoms, Rome, 4.XI.1950. Available here: https://www.echr.coe.int/Documents/Convention_ENG.pdf (Accessed March 2021)

[3] International Covenant On Civil And Political Rights, United Nations, 1967. Available here: https://treaties.un.org/doc/Treaties/1976/03/19760323%2006-17%20AM/Ch_IV_04.pdf (Accessed March 2020)

1.1.1 The Birth of “Artificial Intelligence”

Meanwhile the transparency ideal was flourishing under the international human rights law framework, the technological scenery surrounding “information” was changing drastically. At the time of war, encryption and decryption algorithms were crucial for communicating as well as intercepting the enemy communications. English computer scientist Alan Turing, who was working on deciphering the Enigma code of the Nazi Germany during the war, gave the first lecture on “intelligent machines” in 1947. (McCarthy, 2021) Until then, algorithms were mostly thought as useful for conveying or conserving information. Realizing that algorithmic systems might have the potential to be as good as humans in intellectual tasks, Alan Turing thought about “learning machines” and suggested an imitation game for assessing their success which is known as the “Turing test” (Turing, 1950). Thereby, many researchers started focusing on the utility of the algorithms in decision making processes. In the 1950s Cold War context, reducing intelligence, strategic planning, decision-making and *reason* to algorithmic rules “had spread like wildfire to psychology, economics, political theory, sociology, and even philosophy” which wrongfully replaced the self-critical judgement of *reason* with algorithmic rules of *rationality* such as “the models of game theory, decision theory, artificial intelligence, and military strategy” (Daston, 2013). Going further, American computer scientist John McCarthy invited a group of researchers to discuss “thinking machines” under the name of “Dartmouth Summer Research Project on Artificial Intelligence” in Summer 1956. With this conference, the term “artificial intelligence” (AI) was coined, and AI emerged as a field of research. Although researchers were aware of that algorithmic systems were capable of autonomous decisions making (ADM), the state of the art was still embryonic for the deployment of such systems. Instead, plenty of different decision support systems (DSS) were developed from 1960s on and deployed for helping public administrators and business managers (Power, 2007). Since these systems were rather simplistic tools, their effect on the decision-making processes were not questioned as a matter of transparency at this stage.

The advancement in digitalization and information technologies had to wait for the widespread usage of internet to disrupt the existing legal frameworks and transform the notion of transparency. In the meantime, the post-war optimism yielded to the cold war tension between the socialist Eastern Block and the capitalist Western Block. Thus, war-torn Europe became a strategic field for demonstrating the supremacy of their respective economic models. In the 1970s, neo-liberalism became a transnational governance template in the western block with the help of Organization for Economic Co-operation and Development (OECD), the International Monetary Fund and the World Bank which led to the emergence of fiscal transparency and business-friendliness rankings as well as powerful accounting and audit standard bodies (Mehrpouya & Djelic, 2014). This way, transparency gained a neo-liberal connotation. On the other hand, the conception of transparency among the academic circles and hacker culture of 70s and 80s was rather libertarian than neo-liberal. In fact, the idea of having property rights over algorithms triggered reaction amongst certain software developers. Consequently, computer scientist Richard Stallman founded the “free software movement” in 1983 and gained support from some academics and lawyers. For one, prominent legal scholar Lawrence Lessig has been supporting his approach claiming that sharing the source-codes instead of having copyrights and patents over them is the right way to foster innovation and culture (Lessig, 2020). However, lawmakers in the US and Europe granted intellectual property rights on algorithms and discarded this straight-forward approach to algorithmic transparency (Lessig, 2000).

In 1980, OECD issued the *Guidelines on Trans-border Data Flows and the Protection of Privacy*, considering the increasing level of computer usage for business transactions and international data flow. A year later, Council of Europe developed the first legally binding international instrument concerning personal data protection law; the "Convention 108 for the Protection of Individuals with Regard to Automatic Processing of Personal Data[4]" (Convention 108). This way, data protection law branched out and distinguished itself from the right to privacy. Covering both private and public sectors, Convention 108 became a landmark for regulating digital technologies from a human rights law perspective. Although transparency principle was not explicitly stated in the initial Convention 108, Article 8 granted additional safeguards for the data subjects which were practically requiring transparency. Accordingly, any person had to be enabled "to establish the existence of an automated personal data file, its main purposes, as well as the identity and habitual residence or principal place of business of the controller of the file" and "to obtain at reasonable intervals and without excessive delay or expense confirmation of whether personal data relating to him are stored in the automated data file as well as communication to him of such data in an intelligible form."

With the dissolution of the Eastern Bloc at the end of the 80s, the transparency discourse which played a key role for establishing trust amongst the actors of the post-war world accelerated. Once the integration process of the European States reached a new phase in 1992, the word "transparency" appeared for the first time in a European-level public document. It was used within two different Declarations under the Final Act section of the Treaty on European Union[5] (Maastricht Treaty). One of the declarations ensured "necessary transparency and complementarity" between the emerging "European security and defense identity" and the Atlantic Alliance whereas the other concerned the "right of access to information" with the following paragraph: "The Conference considers that transparency of the decision-making process strengthens the democratic nature of the institutions and the public's confidence in the administration." Remarkably, the transparency conception here is not limited with giving access to documents or information after the decisions are taken but it requires giving access to the "decision-making processes".

1.1.2 The Birth of the Internet

By the 1990s, "world wide web" was invented. Widespread internet connection enabled a huge amount of global data transfers including "peer to peer" sharing which challenged the

[4] Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, Strasbourg, 28.I.1981. Available here: https://rm.coe.int/1680078b37 (Accessed March 2021)

[5] The Treaty on European Union, Signed in Maastricht on 7 February 1992. Available here: https://europa.eu/european-union/sites/europaeu/files/docs/body/treaty_on_european_union_en.pdf (Accessed March 2021)

enforcement of intellectual property rights over software, music, films etc. and encouraged freedom of expression by allowing people to stay anonymous. This new paradigm inspired many political discourses. In 1993, Eric Huges wrote the "Cypher-punk's manifesto" that calls for developing "encryption tools" which are basically computer codes that create "cyphers" for making sure that only the intended receiver of a message can open it. Peculiarly, while demanding privacy for individuals, Huges was loyal to the algorithmic transparency ideal of the free software movement. This was declared in the manifesto as follows: "We publish our code so that our fellow Cypherpunks may practice and play with it. Our code is free for all to use, worldwide." (Hughes, 1993) Another manifesto representing the cyber-libertarian zeitgeist came from John Perry Barlow. In "A Declaration of the Independence of Cyberspace", he claimed that States do not have sovereignty over internet and denounced the 1996 US Telecommunications Reform Act as a treason to the ideals of the founding fathers and Brandeis (Barlow, 1996) In the following year, a historic moment came for the AI researchers. Garry Kasparov, the chess champion, was beaten by an AI system called Deep Blue. At this point, AI systems were already being deployed for various complicated tasks like fraud detection, medical diagnosis and military defense systems.

In response to the rise in digital data transfers and automated data processing with DSSs and ADM systems among businesses and governments, the EU introduced a Directive[6] (hereinafter data protection Directive) in 1995 for protecting individuals in the digital realm which harmonized the European data protection laws and established a legal framework for the ADM systems. Just like the Convention 108, neither transparency nor accountability were explicitly mentioned in the text. Yet, according to Article 12, data subjects had a "right of access" which included the right to obtain "knowledge of the logic involved in any automatic processing of data concerning him at least in the case of the automated decisions referred to in Article 15 (1)". As for the Article 15, the Directive granted data subjects (an interesting way to refer to people) a right to *not to* be subject to solely automated decisions which produce legal or significant effects concerning them, unless the written down criteria are fulfilled. Thereby, a certain amount of transparency regarding algorithmic systems was finally required by law. Thereby, it can be said that the EU took the lead in regulating digital technologies.

As the new millennium begins, the EU enshrined its human rights law framework in the Charter of Fundamental Rights of the European Union[7] (European Charter). Needless to say, the freedom of expression and the right of access to information were also amongst the rights which gained European-level constitutional recognition. Furthermore, Article 11 of the Charter entitled "freedom of expression and information" included respect for freedom and pluralism of the media within this right. Most importantly, the European Charter recognized the right to protection of personal data as a right on its own. Apart from the Article 7 that safeguards protection of privacy, family life and communications, the Charter addressed the

[6] DIRECTIVE 95/46/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. Available here: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046&from=EN (Accessed March 2021)

[7] CHARTER OF FUNDAMENTAL RIGHTS OF THE EUROPEAN UNION (2000/C 364/01). Available here: https://www.europarl.europa.eu/charter/pdf/text_en.pdf (Accessed March 2021)

“protection of personal data” specifically in Article 8 which contains the main principles of data processing as fairness, purpose limitation, consent, legality, right of access and rectification. Thusly, algorithmic transparency became a part of the fundamental rights framework of the European Union in a certain manner.

In 2001, the application of transparency principle and the right of access to information regarding the European Public Authorities became concrete with a European Regulation[8] on public access to the EU documents. Eight years later, Lisbon Treaty gave transparency principle a further constitutional basis as Article 15 of the Treaty on Functioning of the European Union[9] (TFEU) granted ″Any citizen of the Union, and any natural or legal person residing or having its registered office in a Member State, [...] a right of access to documents of the institutions, bodies, offices and agencies of the Union, whatever their medium."

Meanwhile the EU was trying to take steps towards empowering citizens via information, a similar transparency ideal has been gaining popularity on the other side of the Atlantic Ocean as well. In the US, the term “governmental transparency” was established as a fundamental principle that is often used by federal courts in relation to or synonymous to the “open government” which is another popular term that originated there after the second world war (Yu & Robinson, 2012). However, the transparency expectation has been mostly limited with the governmental transparency and it never really covered the source codes of algorithms that governmental bodies were using since the algorithmic systems were often procured from private firms. As for today, although software used by public bodies are considered within “public records” that are disclosable under the federal Freedom of Information Act, there are inconsistencies and unclarities in the application of the rules in the context of algorithms (Fink, 2018). A similar inconsistency exists in the Europe since many recent examples are reported where national authorities did not disclose source codes of algorithms for reasons like trade secrets, intellectual property rights, lack of technical capability or documentation (Algorithm Watch, 2020). Here, it is worth noting that even though if this situation were fixed, that would not suffice to govern societal impacts of the algorithmic systems since they are not only caused by the algorithmic systems used by public authorities. Indeed, the first time that the general public started to wake up to the impact of algorithms in their lives was after the 2008 financial crisis which was engraved by the algorithmic systems that Wall Street firms used[10]. Seven years after the crisis, Frank Pasquale, a legal scholar specialized in algorithms and artificial intelligence, wrote an influential book entitled “black box society” in which he notes that “secret algorithms — obscured by a triple layer of technical complexity, secrecy, and “economic espionage” laws that can land would-be whistle-blowers in prison— still prevent

[8] REGULATION (EC) No 1049/2001 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 30 May 2001regarding public access to European Parliament, Council and Commission documents. Available here: https://www.europarl.europa.eu/RegData/PDF/r1049_en.pdf (Accessed March 2021)

[9]The CONSOLIDATED VERSION OF THE TREATY ON THE FUNCTIONING OF THE EUROPEAN UNION is available here: https://eur-lex.europa.eu/resource.html?uri=cellar:2bf140bf-a3f8-4ab2-b506-fd71826e6da6.0023.02/DOC_2&format=PDF (Accessed March 2021)

[10]See; Sean Dodson, Was software responsible for the financial crisis? Available here : https://www.theguardian.com/technology/2008/oct/16/computing-software-financial-crisis (Accessed March 2021) See also; Richard Dooling, The Rise of the Machines. Available here: https://www.nytimes.com/2008/10/12/opinion/12dooling.html (Accessed March 2021)

us from understanding what is truly going on in many major financial firms." (Pasquale, 2015, p. 103) With this book, Pasquale managed to draw some attention to the fact that opaque algorithms were not only a problem within the financial sector, but they are ubiquitous in the networked society. Just a year later, another remarkable book was written by the mathematician Catherine O'Neil who was working for the finance industry until she was disenchanted and joined the "occupy Wall Street" movement. In her book "Weapons of Math Destruction", she underlined that big data and opaque algorithms were already being deployed for grading the performances of workers; deciding who gets access to credits; who pays higher insurance premiums; who receives which advertisement or information etc. in a way that amplifies inequality and therefore she called for regulation as one of the solutions, including random algorithmic audits by regulators (O'Neil, 2016).

Besides the raising awareness concerning the impact of ADM systems and DSSs, the status quo regarding the reach of transparency conception was also shaken by a series of revelations. Being in the center of it all, the famous non-profit organization called Wikileaks demonstrated that encryption algorithms and internet is not only good for pirating films and music, but it could also enable whistleblowers to leak information to the press while remaining anonymous. Most importantly, the global surveillance disclosures in 2013 by Edward Snowden, an ex-contractor of the National Security Agency (NSA), revealed that the NSA has been tapping into the data acquired by the giant American companies which scaled up in course of time and managed to occupy a huge share of the global web traffic[11]. Namely, the data acquired by Google, Amazon, Facebook, Apple and Microsoft (a.k.a. GAFAM) from their users all over the world were not only used for their commercial interests in an opaque manner but apparently, they were also transferred to the NSA. On top of this, it can be said that the reports on the NSA bugging[12] the EU institutions and wiretapping[13] phone calls of Heads of States including the German Chancellor Angela Merkel heated up the data protection debate in the EU once again and speeded up the renewal of the 20 years old data protection Directive. Consequently, the EU introduced the General Data Protection Regulation[14] (GDPR) in 2016 which came into force on 25 May 2018 and thereby repealed the 1995 data protection Directive.

Being referred as the gold standard of the data protection law, the GDPR triggered the "Brussels effect" and had an international influence by heating up the debate in the US and China besides serving as a model regulation for a lot of States including Japan, Brazil, India, Kenya, South Korea, and California. In terms of algorithmic transparency, the GDPR signifies an important step since the general principles enlisted under the Article 5 includes transparency

[11] See: New York Times, Tech Companies Concede to Surveillance Program. Available here: https://www.nytimes.com/2013/06/08/technology/tech-companies-bristling-concede-to-government-surveillance-efforts.html?ref=technology&_r=0 (Accessed March 2021)

[12] See; Spiegel international, NSA Spied on European Union Offices. Available here: https://www.spiegel.de/international/europe/nsa-spied-on-european-union-offices-a-908590.html (Accessed March 2021)

[13] See; The Guardian, NSA tapped German Chancellery for decades, WikiLeaks claims. https://www.theguardian.com/us-news/2015/jul/08/nsa-tapped-german-chancellery-decades-wikileaks-claims-merkel (Accessed March 2021)

[14] REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

and accountability. Besides explicitly mentioning transparency as a fundamental principle, the GDPR also details the modalities of transparency under Chapter 3 which concerns “rights of the data subject”. On the other hand, Article 22 of the GDPR concerning the automated decisions (which is only a modified version of the provision in the repealed data protection Directive) has a very limited scope. Moreover, the GDPR neither foresees the disclosure of source codes in any situation, nor provides an effective monitoring mechanism to ensure algorithmic oversight. In fact, there is a similar situation with other EU Regulations -like the P2B[15] Regulation- that require transparency. Thus, as it will be explained in the following Chapter, the opacity regarding algorithmic systems continues to a great extent for the time being.

Given the fact that Covid-19 pandemic made online platforms an even more essential part of daily lives of people living in countries with widespread internet connection, addressing the problems like dissemination of false information, hate speech, manipulative, or criminal content as well as unfair economic practices with the help of algorithmic systems is an actual priority for policy makers. As a matter of fact, a debate on algorithmic manipulation was already fuelled before the pandemic with the scandalous reports about the firm Cambridge Analytica which gathered personal data from Facebook to target voters (in many countries around the world including the 2016 U.S. Presidential elections and the Brexit Referendum) in order to manipulate their voting behavior[16]. As the pandemic triggered an “info-demic” and alerted the policy makers, regulating the algorithmic systems that online platforms use became even more crucial. After all, an awareness emerged regarding the need for regulating algorithmic systems in general. Within this context, the EU tries to take the lead once again by proposing prospective regulations which require further algorithmic transparency to address the societal and economic challenges caused by algorithms. Axel Voss -a member of the European Parliaments Special Committee on Artificial Intelligence in a Digital Age- states that the EU is trying to differ from the “big data” approach of the US, and the “big brother” approach of China[17]. Apparently, the EU claims to have a “human centric” approach while Regulating data and algorithms. Considering this endeavour, this thesis focuses on the proposed EU Regulations which promise to push algorithmic transparency further.

1.2 The Problem Statement

Algorithms that affect decision making processes or taking decisions autonomously in numerous ways from curating online content to enabling “artificial intelligence” appliances are

[15] REGULATION (EU) 2019/1150 on promoting fairness and transparency for business users of online intermediation services.

[16] See; Osborne H, Parkinson HJ (2018), “Cambridge Analytica scandal: the biggest revelations so far”. Available here: https://www.theguardian.com/uk-news/2018/mar/22/cambridge-analytica-scandal-thebiggest-revelations-so-far (Accessed March 2021).

[17] See; Eanna Kelly (2020), EU struggles to go from talk to action on artificial intelligence, Available here: https://sciencebusiness.net/news/eu-struggles-go-talk-action-artificial-intelligence (Accessed April 2021)

already widespread. Notwithstanding, opacity is the current norm when it comes to algorithms and most people only learn about them if an insider leaks some information to the press or if journalists conduct "algorithmic accountability reporting" to reveal their impacts by reverse engineering (Diakopoulos, 2014). As the deployment of algorithmic systems proliferates both in private and public sectors, concerns about their impacts grow and calls for transparency become louder. In the European Union, the regulations addressing the use of algorithmic systems require "transparency" in principle. However, defining and establishing transparency is a trickier issue than it sounds since there are legitimate limitations to full transparency and there are different typologies of transparency. (Ananny & Crawford, 2018, p. 976). On top of that, algorithms in use get increasingly complex and become extremely hard to decipher if not impossible. As reported by the US Defense Advanced Research Projects Agency (DARPA), "there is an inherent tension between machine learning performance (predictive accuracy) and transparency; often the highest performing methods (e.g., deep learning) are the least explainable, and the most explainable (e.g., Decision trees) are less accurate." (DARPA, 2016) Therefore, transparency has technical limitations besides the legal side which should be studied well by the policymakers.

In extreme cases where opacity is an inevitable result of technical challenges, the ultimate policy decision is whether to prohibit the deployment of such algorithmic systems or not? But the lines are blurry, and the answer is not always as simple as a yes or a no. Most of the time, the question is: which modality of transparency should be required by law in the given case? This is not only a question of technical feasibility. Besides the technical challenges, there are purely legal puzzles regarding algorithmic transparency. At first glance, one sees a rather clear picture: On the one side, we have trade secrets, intellectual property rights, right to privacy, right to conduct business, concerns about gaming the algorithmic systems or even State secrets if such systems are utilized by public authorities for purposes like national security or mitigating tax fraud. On the other side, we have the pressing need for transparency and oversight since algorithms have serious impacts on human rights, rule of law and functioning of democratic processes. What makes this equation even more complex is that these competing rights do change their sides depending on the situation. For instance, right to privacy might serve as an argument against algorithmic transparency when disclosure of the training data of an algorithmic system is considered whereas it can weight for algorithmic transparency in the context of algorithmic profiling. Indeed, after the Facebook/Cambridge Analytica revelations, European Data Protection Board stated that the use of "sophisticated profiling techniques to monitor and target voters and opinion leaders" raise concerns which surpass a mere privacy and data protection issue and threaten "the trust in the integrity of the democratic process"[18] for underlining the need for transparency.

Another fundamental right enshrined in the European Charter which can weight both for and against algorithmic transparency is the "freedom to conduct a business" that is closely linked to the right to property (FRA, 2015). In general, the freedom to conduct a business has

[18] EDPB, *Statement 2/2019 on the use of personal data in the course of political campaigns* (2019). Available here: https://edpb.europa.eu/sites/edpb/files/files/file1/edpb-2019-03-13-statement-on-elections_en.pdf (accessed March 2021)

a role in justifying intellectual property rights and trade secrets while opposing interventions to the conduct of business. Thus, requiring algorithmic transparency and oversight might mean a derogation for this right. However, the opacity that these rights provide to tech giants might allow them to hamper free market competition by putting competitors in unfair situations, and therefore breach the freedom to conduct a business of others in an indirect way. For instance, such a situation was demonstrated in the European Commission's competition law probes concerning Google search results[19]. Therefore, intellectual property rights, trade secrecy or related freedoms are often balanced against public interests.

Apparently, drafting a legal policy which can safeguard all the rights, freedoms and legitimate interests involved in the different use cases of algorithms is a complicated task. As the state of technology and digital economy changes, the European legal policy tries to catch up. In this endeavor, fine-tuning algorithmic transparency is the key to balance relevant rights against each other if needed and decide which ones overweigh in which situations. With complex algorithms and massive data storage capacities, data driven technologies have become the core of digital economy, digital governance, and even social and political life itself. In this context, the EU introduced the GDPR and the Law Enforcement Directive[20] back in 2016 both of which came into force in 2018. This way, EU regulated the whole lifecycle of personal data including the processing activities by automated means. Yet, a satisfying level of algorithmic transparency was not harbored in the European data protection laws.

Besides the general transparency principle in Article 5 (a), the GDPR harbors data subjects right like "the right of access", obliges data processors to give meaningful information regarding the data processing as well as prohibiting fully automated decisions unless certain conditions and due process rights are met, and a limited explanation of the automated decisions is provided. As the wording of the Article 22 (1) of GDPR expresses that the conditional prohibition of automated decisions applies for "…a decision based solely on automated processing…", the DSSs are not covered under this article. Therefore, the due process rights and transparency requirement regarding algorithmic systems do not apply to algorithmic systems as long as there is a "human in the loop". Most importantly, the obligations of data controllers and the rights of data subjects remains bounded with intellectual property rights and trade secrets. For instance, Recital 63 states that the right of access in the GDPR should not adversely affect "trade secrets or intellectual property and in particular the copyright protecting the software". By the same token, Articles 23(1), 89 (2) and 89 (3) of the GDPR allows Union or Member State law to limit the obligations and rights under Articles 12-22 in some circumstances where there is public interest.

The shortcomings of the GDPR in terms of algorithmic transparency is not only due to its substance. There is also an enforcement problem which renders the Regulation rather toothless. This is because the data protection authorities (DPAs) were not given a systematic and structural oversight role regarding algorithms in practice, although they theoretically had

[19] For the full story, see; Kristie Pladson (2020), Companies plead with EU regulators for action on Google. Available here: https://www.dw.com/en/google-antitrust-eu-european-union/a-55623857 (Accessed March 2021)

[20] DIRECTIVE 2016/680 with regard to data processing for law enforcement purposes which repealed Council Framework Decision 2008/977/JHA.

the competence to ensure that personal data processing is in accordance with the principles enshrined in the GDPR. Moreover, the DPAs are not given sufficient human, financial and technical resources[21]

Despite the shortcomings, the EU took the first step in regulating digital technologies by setting the standards for "data" which is the first component of the digital life. As for the algorithms, the EU seems to prefer addressing different angles of the issue with different regulations. For one, the so called P2B Regulation requires the online intermediation services to be transparent and fair in order to safeguard the "business users" of such services and "corporate website users" in relation to search engines. Yet again, the transparency requirement in this Regulation does not require the disclosure of algorithms according to the Article 5 (6) and it is limited with the trade secrets -just like the GDPR- as the Article notes that transparency regarding the ranking systems shall be without prejudice to the EU Trade Secrets Directive[22]. Following Recital 27 of the P2B Regulation, although online intermediation service providers are not required to disclose their algorithms, they are obliged to provide the main parameters determining the ranking system and a description "based on actual data on the relevance of the ranking parameters used." Similarly, Article 7(4a) of the EU Directive on consumer rights as amended by Directive 2019/2161[23] requires that consumers must be informed of the "main parameters determining the ranking of products presented to the consumer as a result of the search query and the relative importance of those parameters, as opposed to other parameters." Here, it is worth noting that different regulations address different fields of law which contain different actors. Therefore, different regulations need to exemplify different modalities of transparency which are adapted to the different contexts. The common ground of these different transparency modalities is that they all protect the source codes of algorithms in use from being exposed. However, this amount of transparency might not always suffice in an age where the lack of control over intermediary platforms and the algorithmic systems they use can have anti-competitive, manipulative, polarizing or discriminatory effects. Furthermore, this lack of oversight can infringe various categories of human rights from right to privacy, right to data protection, right to freedom of speech to causing indirect harms to right to bodily integrity, right to live and so on.

[21] This issue is also stressed by the European Commission and the European Parliament. See; COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL: Data protection as a pillar of citizens' empowerment and the EU's approach to the digital transition - two years of application of the General Data Protection Regulation. COM (2020) 264 final of 24.6.2020.

[22] DIRECTIVE (EU) 2016/943 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 8 June 2016 on the protection of undisclosed know-how and business information (trade secrets) against their unlawful acquisition, use and disclosure.

[23] DIRECTIVE (EU) 2019/2161 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 November 2019 amending Council Directive 93/13/EEC and Directives 98/6/EC, 2005/29/EC and 2011/83/EU of the European Parliament and of the Council as regards the better enforcement and modernisation of Union consumer protection rules

In February 2020, the European Commission released a communication on its Data Strategy[24] and a White Paper on Artificial Intelligence[25], fleshing out the perspective of the EU on regulating the digital technologies. To begin with, the Commission proposed a Data Governance Act[26] (DGA) and called for an AI Regulation as part of the set of measures announced in the strategy. Later on, the Commission proposed the Digital Services Act[27] (DSA) and Digital Markets Act[28](DMA) which include new obligations for the online platforms. Finally, the Commission presented the Artificial Intelligence Act (AI Act)[29] in March 2021 concretizing the roadmap drafted in the White Paper on AI. With these prospective Regulations, the EU wants to complement its legal framework covering the two main elements composing "Artificial Intelligence" (AI) as well as other algorithmic systems: *data* and *algorithms*.

Even though the DSA, the DMA and the AI Act have different subject matters, these prospective regulations are meant to complement each other as well as the other relevant regulations that are already in force and serve similar policy goals. Remarkably, the fundamental principles enshrined in the relevant regulations including the DSA, the DMA, the AI Act and the GDPR converge around transparency, fairness, and accountability ideals. Although the general principles in these regulations are all interrelated and equally valuable, it can be said that the transparency principle plays a slightly different role. In fact, all the principles in these regulations are dependent on "transparency" in one sense since getting access to the information regarding the algorithms at stake is necessary for having human oversight, establishing accountability, or assessing fairness, safety etc. In other words, one needs to know how an algorithmic system works in detail to assess whether it meets the criteria laid down by the regulations. Thus, it can be said that the transparency principle is a constituent of the other principles and recalibrating it is the most crucial point.

1.3 The Research Question

The European discourse on regulating new technologies suggests that the EU differs from other global powers with a "human centric" approach that would safeguards rights and

[24] Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. A European strategy for data, COM (2020) 66 final of 19.2. 2020.

[25] WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust, COM (2020) 65 final of 19.2.2020.

[26] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on European data governance (Data Governance Act), COM (2020) 767 final of 25.11. 2020.

[27] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC, COM (2020) 825 final 2020/0361 (COD) of 15.12.2020.

[28] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on contestable and fair markets in the digital sector (Digital Markets Act), COM (2020) 842 final 2020/0374 (COD) of 15.12.2020.

[29] Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL LAYING DOWN HARMONISED RULES ON ARTIFICIAL INTELLIGENCE (ARTIFICIAL INTELLIGENCE ACT) AND AMENDING CERTAIN UNION LEGISLATIVE ACTS, COM (2021) 206 final of 21.4.2021.

freedoms. Developing such a legal framework for algorithmic systems and taking the lead in the field of regulating digital technologies such as artificial intelligence is an ambitious goal indeed. As the European legislators discuss the abovementioned prospective Regulations, this thesis reflects on whether the EU delivers its promise by striking the right balance between the competing rights within its legal framework. In particular, “algorithmic transparency” is held under the microscope since transparency emerges as the prominent tool for regulating algorithms. Thereby, this thesis questions; why is transparency needed in the context of algorithmic systems? What does transparency mean in the context of algorithmic systems? What are the obstacles for it? And aims to humbly contribute to the legislative efforts by evaluating the initial drafts of the DSA and the AI Act, drawing lessons from the GDPR experience. Therefore, how would these new regulations recalibrate algorithmic transparency in the EU and what else could be done are the focal questions for this thesis.

1.4 Interdisciplinary State of Knowledge

Algorithmic transparency is a multifaceted issue. Therefore, the literature review for this thesis covers many different disciplines which are necessary to assess the implications of algorithmic transparency (or the lack thereof). Besides legal doctrines, case law and official documents issued by governmental or regulatory bodies; the pieces written by legal scholars, AI ethicists, AI developers, journalists, sociologists, computer scientists, philosophers, historians, political scientists, activists, data scientists and mathematicians were especially helpful for better grasping the issue at stake. Since the relevant literature is available in the Bibliography, this sub-chapter only contains a brief sample of the basic academic consensus that this thesis is built upon.

For starters, there exists a growing amount of literature and official documents on algorithmic transparency since the issue is not new. First and foremost, computer algorithms themselves have been thought as a new type of regulation that was born with the cyberspace (Lessig, 1999). In course of the time passed since Lessig first wrote “code is law”, algorithmic systems became more and more complex and widespread. Notwithstanding, opacity regarding them grew due to intellectual property rights, trade secrecy or other concerns such as technical complexity. Not that long ago, the well-known books of Mayer-Schönberger and Cukier (2013), Frank Pascal (2015) and Catherine O’Neal (2016) among other books as well as many academic papers made the “black box algorithms” issue more visible and stressed the need for regulation[30]. As for today, “technological solutionism” (Morozov, 2013) and “techno-chauvinism” (Broussard, 2018) are refuted many times and algorithmic systems are no longer seen as neutral or objective agents capable of solving every issue without any need for oversight or transparency. On the contrary, it is a well-established fact that they are value-laden and fallible[31]. To establish accountability and fairness, solutions like “algorithmic audits” have been on the table for some time. Lastly, it has been suggested to establish a pre-market approval body for algorithms, drawing inspiration from the U.S. Food and Drug administration (Tutt, 2017).

After the EU introduced the GDPR which covers “automated data processing”, further interesting literature on algorithmic transparency emerged. In particular, the “right to explanation” debate[32] helped algorithmic transparency to become a priority for the European Law. On the other hand, some scholars like Edwards and Veale thought that the search for a right to an explanation was a new kind of fallacy just like “meaningless consent.”   (Edwards & Veale, 2017, p. 81) In the meanwhile, scholars Wachter, Mittelstadt and Floridi brought up the “algorithmic audits” once again in their paper concerning the shortcomings of the GDPR in terms of applying transparency and accountability principles to automated decision-making (Wachter, et al., 2017, p. 46). Furthermore, Wachter and Mittelstadt suggested a “right to reasonable inferences” that would provide individuals with the information that enables them to practice their rights under the GDPR (Mittelstadt & Wachter, 2019)

For the time being, the EU works on regulating algorithmic systems with separate regulations. Namely, the DSA, the DMA and the AI Act will introduce some obligations including targeted transparency policies, third party audits and certification requirements.

[30] See; inter alia. Kate Crawford (2016), Can an Algorithm be Agonistic? Ten Scenes from Life in Calculated Publics, 41 SCI. TECH. & HUM. VALUES 77; David Beer (2016), The Social Power of Algorithms, 20 J. INFO. COMM. & SOC. 1; Rob Kitchin (2016), Thinking Critically About and Researching Algorithms, 20 INFO. COMM. & SOC’Y 14; Michael Ananny (2015), Toward an Ethics of Algorithms: Convening, Observation, Probability, and Timeliness, 41 SCI. TECH. & HUM. VALUES 93; Nicholas Diakopoulos (2015), Algorithmic Accountability: Journalistic Investigation of Computational Power Structures, 3 DIGITAL JOURNALISM 398; Christian Sandvig (2015), Seeing the Sort: The Aesthetic and Industrial Defense of “The Algorithm”, 12 J. NEW MEDIA CAUCUS, 1; Zeynep Tufekci (2015), Algorithmic Harms Beyond Facebook and Google: Emergent Challenges of Computational Agency, 13 J. ON TELECOM. & HIGH TECH. L. 203; Malte Ziewitz (2015), Governing Algorithms: Myth, Mess, and Methods, 41 SCI. TECH. & HUM. VALUES 3; Taina Bucher (2014), ‘Want To Be on the Top?’ Algorithmic Power and the Threat of Invisibility on Facebook, 14 NEW MEDIA & SOC’Y 1164; Jenna Burrell (2016), How the Machine ‘Thinks’: Understanding Opacity in Machine Learning Algorithms, 3 BIG DATA & SOC’Y 1; Danielle Citron & Frank Pasquale (2014), The Scored Society: Due Process for Automated Predictions, 89 WASH. L. REV. 1.

[31] See; inter alia. Brent Daniel Mittelstand, Patrick Allo, MariarosariaTaddeo, Sandra Wachter, Luciano Floridi (2016), The ethics of algorithms: Mapping the debate, in Big Data & Society; Lisa Gitelman (editor) (2013), “Raw Data” is an Oxymoron; Danah Boyd, Kate Crawford (2012), Critical questions for Big Data: provocations for a cultural, technological, and scholarly phenomenon, Information, Communication & Society, 662.

[32] See; Inter alia. Bryan Casey, Ashkon Farhangi & Roland Vogl (2019), Rethinking Explanable Machines: the GDPR’s “Right to Explanation” Debate and the Rise of Algorithmic Audits in Enterprise, Berkeley Tech. Law Journal; Margot E. Kaminski (2019), The Right to Explanation, Explained. Berkeley Tech. Law Journal.

Since the Commission introduced the proposals recently, there are not many resources available directly addressing these prospective regulations for the time being.

1.5 Methodology, Data and Resources

This thesis is a result of a library-based research. As it concerns the legislative efforts of the European Union, the official documents such as white papers, scientific reports, communications, and drafts of the proposed regulations are its primary resources. For the academic literature, the online resources that Sciences Po Bibliothèque grants access to as well as open access research papers and books have been used besides the personal resources of the author. Notably, although the measures taken in the context of the Covid-19 pandemic prevented the author from conducting research by visiting the libraries physically, this was not a limitation for the purposes of this thesis due to the fact that the relevant material for the subject of this thesis is available -sometimes exclusively- online. The quality of the data and resources were ensured by giving weight to the expertise and verification. At a minimum, the primary resources were taken as the official documents whereas the academic resources that are peer-reviewed or well acknowledged were given priority.

In general, the thesis has a multi-disciplinary approach which revolves around a qualitative analysis on the relevant legal doctrines, technological challenges as well as public policy concerns. The initial research for the thesis was conducted on the history of algorithmic systems as a new technology and the way that they were entreated by the lawmakers. This “back to the basics” approach revealed that the secrecy regarding algorithmic systems was amplified by legislative policy decisions rather than the technical complexity. Although the thesis focuses on the EU Law, this approach adds a pinch of comparative aspect to it since the origins of the technology at stake were traced back to the UK and the US. Needless to say, there was already a comparative law aspect as the Member States have different legal frameworks or legal practices. Given that the EU is trying to take the lead in regulating digital technologies, the thesis contains a qualitative assessment of some EU regulations in force as well as the initial drafts of the proposed DSA and the AI Act.

CHAPTER 2: ALGORITHMIC TRANSPARENCY

This chapter aims to provide an analysis regarding the conception of “algorithmic transparency.” The first sub-chapter contains a brief inquiry on the reasons of transparency demand for algorithmic systems since understanding the problems that are caused or amplified by opacity is the first step for exploring the transparency modalities that could tackle them. The second sub-chapter deals with which kind of information could be provided concerning algorithmic systems. Given that there are many ways to create algorithmic systems which all have different constituents and phases, the chapter indicates that a meticulous approach is needed for achieving transparency in such a technically complex and heterogenous field. Then, the next sub-chapters elaborate the technical and legal reasons of opacity by examining the legal framework of the EU in particular. Finally, the concluding sub-chapter reflects on how transparency could be tailored for the algorithmic systems. Thereby, the Chapter remarks that the word “transparency” is often used in a narrow sense that signifies the availability of limited information for the individuals concerned or the general public, although further transparency towards other subjects like the public authorities can also be of crucial importance in the context of algorithmic systems.

2.1 Why Transparency?

As Brownsword writes, “it is axiomatic that good (democratic) governance entails transparent governance.” (2005, p. 15) Indeed, “open government” and “transparency” are often seen as qualities of a modern and democratic State. The OECD recognizes open government as “a culture of governance that promotes the principles of transparency, integrity, accountability and stakeholder participation in support of democracy and inclusive growth” which is “critical to building citizen trust” (OECD, 2017). Although there is such a consensus on the transparency principle among the so-called democratic States, they often struggle when it comes to applying it, especially in the context of algorithmic systems. Besides the technical challenges, this is also due to the different policy decisions which can translate into giving weight to different interests and balancing competing human rights. As for today, a neo-liberal paradigm seems to be dominant whose transparency demand mostly concerns financial transparency from nation States and inclusion of the business world in the decision-making processes. This approach supports outsourcing public tasks to private sector (sometimes under the pretext of digitalization) and yet opposes algorithmic transparency. However, this status quo seems to change as the negative impacts of algorithmic opacity are being demonstrated again and again.

As a matter of fact, the demand for algorithmic transparency has many solid reasonings. Let alone the concerns about polarization, unfairness, discrimination, accountability, manipulation, disinformation and so on, transparency regarding algorithms is needed most of all for the “rule of law.” In this regard, Hildebrandt reminds that “law” differs from administration or discipline because of the opportunity to “challenge the rules validity in view of higher legal norms and to contest its application in particular instances" (2013) by drawing inspiration from Brownsword's thoughts on techno-regulation (2005). In the absence of transparency, algorithmic systems regulate the society without giving individuals a chance to challenge imposed rules. Furthermore, as private vendors get involved in algorithmic design which plays a role in e.g., the governance of a city, new risks including "opacity, public disempowerment, and loss of accountability" emerge because public officials may not "participate in and may be unaware of the policy decisions that are incorporated into those algorithms." (Brauneis & Goodman, 2018, p. 116) Thereby, lack of transparency means an unjustified transfer of power from public to private sector. This way, decisions concerning the public are being taken behind closed doors and being implemented in a secretive way. For instance, programmers need to decide on whether to favour one kind of error over another or treat them the same while developing DSSs or ADM systems. In the context of criminal law, the difference between giving weight to a "false positive" or a "false negative" translates into a decision concerning the fundamental principles of the legal system like whether to risk falsely imprisoning an innocent person or wrongly releasing a guilty person. Therefore, the public needs to be sure that programmers are following the existing legal principles and the policy decisions are taken in a democratic and transparent way.

The most-studied problems with opaque algorithmic systems are fairness and accountability. Especially, unjustified discrimination based on legally protected characteristics of individuals like race or gender is a typical algorithmic fairness concern[33]. Indeed, since human judgment can be involved in designing the system, pre-classifying training data and adjusting thresholds and parameters, the algorithms can inherit discrimination (Burrell, 2016, p. 3). Moreover, the root of the bias in an algorithmic system can also be at the machine learning process (Kuner, et al., 2017, p. 1). For instance, a system can have bias because of the selection of training data which underrepresents some groups. Consequently, systems can include people to wrong profiles, be self- reinforcing and unjustly discriminate.

In the US, the algorithmic system called COMPAS constituted a clear example for potential negative impacts of algorithmic systems. In the Loomis Case[34], Supreme Court of Wisconsin upheld the use of COMPAS algorithm that is used in Courts to inform judges with "recidivism scores" for taking probation decisions. Yet, judges do not have access to the internal working model of the algorithmic system which is protected as a trade secret. Investigative journalists demonstrated that this algorithm was discriminatory against black defendants although it did not include such an information as an input[35]. In other words, "ban the box" policy in the questionnaire did not overcome discrimination since other inputs worked as proxies to racial profiling[36]. This example illustrates how opacity regarding algorithmic

[33] See Chapter 1.

[34] State of Wisconsin v. Eric L. Loomis (2016). Available here: https://www.wicourts.gov/sc/opinion/DisplayDocument.pdf?content=pdf&seqNo=171690 (Accessed March 2021)

[35] Jeff Larson et al, How We Analysed the COMPAS Recidivism Algorithm. Available here: https://www.propublica.org/article/how-we-analyzed-the-compas-recidivism-algorithm (Accessed March 2021)

[36]The questionnaire for the COMPAS algorithm is available here: https://www.documentcloud.org/documents/2702103-Sample-Risk-Assessment-COMPAS-CORE.html (Accessed April 2021)

systems can automate, scale up and hide unfair discrimination. As a matter of fact, most of the algorithmic decisions (or informative outputs) regarding persons are based on algorithmic profiling. In brief, profiling algorithms match the past choices of the concerned individual with the past choices of other individuals that have the same characteristics in order to predict their behaviour. Therefore, algorithmic profiling is a "tyranny of the past" as well as a "tyranny of the others' choices." (G'sell, 2019) To avoid this unfair situation, algorithmic systems should be transparent enough to ensure that there is a meaningful human oversight.

Another significant example for the potential risks of public procurement of opaque algorithmic systems from private sector is the "smart cities" movement. This business model locks-in the city services to certain private vendors and makes public administration to lose control over the management of public infrastructure (Meyer, 2017). What is more, the algorithms do not necessarily have to be deployed by the public sector in order to raise such public policy tensions. A branch of the literature on social impact of algorithms has been drawing inspiration from the "actor-network theory" which focuses on the entanglement of human bodies and thoughts with non-human objects like devices in socio-technical networks[37].With such an approach, Dominique Cardon and Maxime Crépel isolate algorithmic systems as "calculating agents" which are not only shaped by the programmers but also by the "data produced by users, the behaviour of other actors or regulatory principles imposed by external sources." (2019, p. 7) In their study, they frame algorithmic agency as a "public problem" by constituting a repertoire of cases that enable the identification of the different facets of problematic situations[38]. Notably, their repertoire consists of algorithmic systems utilized by private firms like Uber and Airbnb which also have a significant impact on "territorial regulation of the city" without being used by the public sector. A recent example that demonstrates the need for transparency regarding the algorithmic systems utilized by the private sector is the propagation of fake news via social media platforms during the Covid-19 pandemic. For one, Facebook has been taking -good or bad- policy decisions by itself about this issue which has an impact on the fight against the pandemic and therefore on the public health as well as the right to life of certain individuals[39]. Yet, the role of the algorithmic systems on the info-demic remains as a secret. The "echo chambers" in the social media and streaming platforms certainly have a role in radicalization, polarization, and dissemination of fake news although Nick Cleg, the vice president for the global affairs of Facebook, puts the blame -at least partially- on the people for such negative impacts (Cleg, 2021). After all, it is possible to list a sheer number of different examples of undesirable situations amplified or enabled by the algorithmic systems that private sector deploys.

[37] See; Latour, B. (2005) Reassembling the Social: An Introduction to Actor-network-theory. Oxford: Oxford University Press.

[38] Here, "public problem" is a translation as this text is translated by Renuka George from the original version in French. Originally, the authors use the word "problème public" refferring to the following article: Daniel Cefai, "La construction des problèmes publics. Définition de situations dans des arènes publiques", Réseaux, 1996, n°75, pp. 43-66.

[39] For a report on the policy decisions of FB concerning the Covid-19 fake news issue see; Héloïse Théro Emmanuel Vincent, Investigating Facebook's policies to tackle misinformation during the COVID-19 pandemic. Available here: https://medialab.sciencespo.fr/productions/2020-investigating-facebooks-policies-to-tackle-misinformation-during-the-covid-19-pandemic-heloise-thero-emmanuel/ (Accessed April 2021)

Eventually, all the algorithmic systems that have serious societal impacts should be transparent and open to oversight as well as public scrutiny in one way or another. Unless a meaningful level of transparency is achieved, establishing accountability, or assessing an algorithmic systems relationship to governance, fairness and politics as well as its safety and performance would not be possible.

2.2 Transparency and Algorithms: neither a mismatch nor a concord

In order to understand what transparency means in the context of algorithms, it might be useful to define what is an algorithm first. As defined pretty neatly in a study that was commissioned by the "Panel for the Future of Science and Technology" of the European Parliamentary Research Service, an algorithm is "an unambiguous procedure to solve a problem or a class of problems. It is typically composed of a set of instructions or rules that take some input data and return outputs. As an example, a sorting algorithm can take a list of numbers and proceed iteratively, first extracting the largest element of the list, then the largest element of the rest of the list, and so on, until the list is empty. Algorithms can be combined to develop more complex systems, such as web services or autonomous cars. An algorithm can be hand-coded, by a programmer, or generated automatically from data, as in machine learning." (Castelluccia & Métayer, 2019)

Indeed, "algorithm" is a very broad concept which covers *Boeuf Bourguignon* recipe as well as computer algorithms like Google's PageRank. However, the word "algorithm" has a more specific meaning today since no one would virtually call food recipes as algorithms in the kitchen. People often refer to the "algorithms" indicating the ones that are implemented in the form of a computer program, software, or information system. In order to be executed by computers, algorithms must be expressed in a specific programming language like Phyton, C, R, Java etc. that enables humans to read/write the "source code" from which a binary code that consists of ones and zeros is generated that is called the "object code." Furthermore, the word "algorithm" is also often used interchangeably with "algorithmic system" which is basically a system that is created by combining different algorithms. With the rise of "big data" analysis, algorithmic systems that infer information from data in order to achieve a given task proliferated and became focal for the algorithmic transparency debate given the intrinsic opacity of certain types of them. As for today, the prominent methods for developing such complex algorithmic systems known as "black box algorithms" are machine learning (ML) and deep neural networks (DNN). With ML processes, algorithmic systems can test the strength of correlations between many variables and outcomes for creating a "model" that can estimate the likelihood of future behaviour or events (output) when given relevant facts as inputs (Hill, 2016). Such algorithmic systems are increasingly deployed for automated decision making or supporting decision making processes. Simply, ML processes have different phases that could be disclosed such as data collection, data preparation, choosing/designing a model, training the model, evaluation, parameter tuning and the prediction. For instance, the algorithmic systems are often built on a relatively small variety of analytic techniques to discover correlations some of which are support vector machines, linear regressions, logistic regressions, polynomial regressions, neural networks, and random forests (Shalev-Schwartz & Ben-David, 2014). Therefore, documenting which techniques are chosen and why is possible. Furthermore, the policy choices like the "feature engineering" decisions regarding the importance of certain kinds of errors (weighing of accuracy, precision, and recall) or decisions like excluding otherwise relevant data for a reason could be documented and explained. Yet, it is worth noting that simply disclosing information regarding these phases do not necessarily establish "transparency." In fact, it could even be harmful. For instance, disclosing training data without any redaction could damage the right to data protection and privacy of the people whose personal data is used for training a ML model.

In light of this clarification regarding the term "algorithm", it might be useful to have a look at another study commissioned by the same unit of the European Parliament which elaborates transparency in the context of ADM systems by stating that: "Depending on the type and use of an algorithmic decision system, the desire for algorithmic transparency may refer to one, or more of the following aspects: code, logic, model, goals (e.g. optimisation targets), decision variables, or some other aspect that is considered to provide insight into the way the algorithm performs. Algorithmic system transparency can be global, seeking insight into the system behaviour for any kind of input, or local, seeking to explain a specific input - output relationship." (Koene, 2019) As it can be also observed by this study, transparency may become something different than simply disclosing raw data in the context of the algorithmic systems. Definitely, for establishing transparency in the context of algorithmic systems, the first question to ask is which kind of information should be provided, starting with whether complete transparency is necessary or partial transparency and interpretability is sufficient? Moreover, transparency should be specified by indicating who should be given information and when.

2.3 Reasons of Opacity

There are many reasons of algorithmic opacity, some of which are organizational or technical challenges, whereas some are legal obstacles. For instance, algorithmic systems can be kept secret due to the concerns about "gaming the system". Not surprisingly, neither public authorities nor private firms prefers to play by "revealing their hand" if the game is somehow serious. As a matter of fact, secrecy can serve the "public interest" sometimes. Not revealing an algorithmic system concerning cyber-security, or detecting financial fraud is simply logical. Disclosing such an algorithmic system to the general public would allow malevolent actors to game the system maybe by training a counter-algorithm that would enable them to avoid scrutiny. In machine learning parlance, this kind of a process is called "adversarial learning", and this is why such algorithms are often dynamic and opaque. In a similar fashion, firms whose competitive advantage might be affected by the disclosure of its algorithmic systems might not share the documents containing the relevant information with e.g., States outsourcing their algorithmic systems or with the individuals requesting information. Notably, this kind of private interests can be in conflict with "public interest" sometimes. On the other hand, opacity can stem from unintentional reasons as well. First of all, "explaining" the logic of an algorithmic system might be an intrinsically problematic task depending on the methods used. Furthermore, opacity can stem from pretty simple issues like the lack of documentation. In the absence of specific documentation requirements, firms or governments might not organize themselves to be able to provide enough information when requested.

In order to establish an algorithmic transparency regime that serves the public interest, it is necessary to grasp the reasons of opacity. Therefore, the following sub-chapters concern the main technical and legal obstacles for algorithmic transparency.

2.3.1 Technical Opacity

Lack of transparency regarding algorithmic systems can be a technical concern sometimes. In fact, understanding an algorithmic systems decision-making process is not easy even if the source codes are disclosed. Such a transparency would only make sense if the person getting access to the algorithms understands their language[40]. To be meaningfully informed about the process, the general public needs experts, or even other algorithmic systems to "translate" the algorithmic decision-processes into words corresponding to human reasoning. As Jenna Burrell articulates perfectly, opacity can stem from the "mismatch between mathematical optimization in high-dimensionality characteristics of machine learning and the demand of human-scale reasoning and semantic interpretation." (2016, p. 2) In other words, making sense out of algorithmic transparency is a very delicate issue since a lot could be lost in translation. In this respect, the algorithmic systems based on machine learning or deep neural networks -which are often called as "artificial intelligence"- pose a great technical challenge. The systems based on machine learning (ML) or deep neural nets (DNN) are not transparent about their "reasoning" except giving some information about "weights" or statistical correlations of the data items, unlike the simpler systems -with decision trees or similar structures- whose "reasoning" is more understandable (Burrell, 2016). Thus, the first challenge for the transparency ideal is the complexity of the "models" created after the training processes of systems based on ML and/or DNN since they are based on mathematical correlations instead of logical causations. In particular, explaining the decisions of DNN systems by the "weights" given in a multilayer neural net seems like a very hard task although there are ongoing research projects for training algorithms which quantify the influence of the input variables on the outputs (Goodman & Flaxman, 2016, p. 6).

Another reason of opacity regarding algorithmic systems can be the existence of a dynamic model. As a matter of fact, internal decision logic of algorithmic systems based on ML can be altered as they keep "learning" (Burrell, 2016, p. 5). Simply put, some algorithmic systems might change in time and thus only an ex-post intervention for opening the "black box" could provide something close to an explanation about a given event. Using this kind of

[40] This is not merely an issue of knowing the relevant programming language since the decision processes of complex algorithms are also dependent on other variables like the training method and the data utilized.

dynamic models are preferred in tasks that require adaptation such as monitoring credit card transactions in order to reveal fraud (Datatilsynet, 2018, p. 10).

In brief, when the opacity of an algorithmic system is due to technical complexity, it is only partially interpretable, and if its model changes in time, it needs to update the explanations each time the model changes. In fact, "explainable artificial intelligence" endeavours often have interpretability constrains and their accuracy tends to be lower than "black box" models. Still, there are promising methods like developing explanation interfaces like saliency visualisation, heat maps or natural language explanations for giving significant insights to the rationale of an algorithmic system (Gunning & Aha, 2019). Lastly, since an "explanation" gains its meaning within the given context and from the perspective of the receiver, a one-size-fits-all explanation would not be possible. For instance, from the perspective of the persons affected or end-users, "counterfactual explanations" might be more useful than opening the "black box" (Wachter, et al., 2018) Moreover, if the person is a minor, a simpler explanation would be needed. After all, "transparency is not an end in itself, but an interim step on the road to intelligibility." (Pasquale, 2015, p. 8)

2.3.2 Legal Opacity

The prominent legal reasons of algorithmic opacity are intellectual property (IP) rights. In fact, intellectual property seems to "create more problems than solutions" in the context of algorithmic systems since the limited exceptions to them only grant "quasi-rights" that hardly establishes transparency (Diega, 2018). In the EU, computer programs are subject to a specific copyright protection regime which protects the expression of a computer program (not the ideas, methods and principles which underlie any element of it). According to the Article 8 (2) and the Recital 16 of the Software Directive[41], there is a right to "observe, study or test the functioning of the program in order to determine the ideas and principles which underlie any element of the program." On the other hand, the copyright protection under the Article 10 of TRIPS[42] agreement covers the source code, object code as well as the relevant data set if it constitutes an intellectual creation. Accordingly, copyright protection can create a legal "black box" that is hard to penetrate. Similarly, another legal ground for the opacity regarding algorithmic systems could be a patent protection. In fact, algorithmic systems are regarded as mathematical methods within the meaning of the European Patent Convention[43]. Therefore, they are not patentable as such except if they are used as part of an AI system that contributes to producing a further technical effect that can fall under the Article 52(3). Yet, an increasing number of AI-related patents are being granted[44]. Although patents are less of a problem by

[41] DIRECTIVE 2009/24/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 23 April 2009 on the legal protection of computer programs

[42] See; The Agreement on Trade-Related Aspects of Intellectual Property Rights of the World Trade Organization.

[43] Convention on the Grant of European Patents (European Patent Convention) of 5 October 1973 as revised by the Act revising Article 63 EPC of 17 December 1991 and the Act revising the EPC of 29 November 2000.

[44]See; European Parliament resolution of 20 October 2020 on intellectual property rights for the development of artificial intelligence technologies (2020/2015(INI)).

having shorter time limits, they do veil the algorithmic systems as well. However, the biggest legal obstacle for the algorithmic transparency is the concept of "trade secret" and its aggressive use. For instance, the patent[45] of Google's PageRank was expired in 2019 whereas the rest of the search engines algorithmic system remains as a trade secret.

In 2016, both the EU and the US updated their legal frameworks concerning trade secrets to provide more legal clarity. The US federalized its trade secrets law with the Defend Trade Secrets Act[46] which allows trade secret holders to seek remedy in federal courts in certain situations whereas the EU adopted Directive (EU) 2016/943 on the protection of undisclosed know how and business information (trade secrets) against their unlawful acquisition, use and disclosure (hereinafter, "Trade Secrets Directive"). From a legal standpoint, fundamental rights cannot be overruled by a directive. However, EU Directives can harmonize national laws by defining the criteria under which the judges should balance competing rights against each other. Indeed, the EU Trade Secrets Directive demarks the right of businesses to the secrecy of their information which has commercial significance as opposed to the fundamental rights of individuals such as "right to information" enshrined in the European Charter and other constitutional documents. The notion of trade secret as set forth by Article 2 of the EU's Trade Secrets Directive refers to "information", that is (a) secret; (b) has commercial value due to its secrecy; and (c) has been subject to reasonable steps to keep it secret. Thus, algorithms are perfectly capable of fitting in this definition. In fact, if an algorithm is not published online, inscribed in a public register, or presented in a public event like a trade convention, it is automatically kept secret, and most of the time, it has an economic value due to its secrecy. Thereby, information regarding them is often exempted from "the right to information" of the individuals.

Another concept that the algorithms could fall under is "know how." As stated in the very first Recital of the Trade Secrets Directive, "valuable know-how and business information, that is undisclosed and intended to remain confidential, is referred to as a trade secret". Indeed, algorithms can be considered as "know-how" being pieces of information that are functional to production or distribution of a commodity as well as instructions on how to pursue a task. Article 4(2)(a) concerning the unlawful acquisition of trade secrets refers to "any documents, objects, materials, substances or electronic files containing the trade secret or from which the trade secret can be deduced." In the context of algorithms, an element that allows the deduction of a trade secret can be the missing information for reverse engineering an algorithmic system since complex algorithmic orders cannot be simply revealed by the personalized services or products that they contribute or create.

When an algorithm that is protected under an IP law regime is disclosed to a third party, the third party is usually bound to keeping it secret under a confidentiality agreement or restrictive contractual clauses. For instance, firms may only licence the object code of the algorithms and keep the source code secret as well as banning the reverse engineering of the

[45] The Patent of the PageRank algorithm is available here: https://patentimages.storage.googleapis.com/db/8f/cb/dad63e985797ec/US7058628.pdf (Accessed April 2021)

[46] PUBLIC LAW 114–153—MAY 11, 2016 DEFEND TRADE SECRETS ACT OF 2016.Available here: https://congress.gov/114/plaws/publ153/PLAW-114publ153.pdf (Accessed April 2021)

object code via a specific contractual clause as it often happens with ordinary pieces of software (Maggiolino, 2019, p. 8). In this regard, Article 3(1) of the Trade Secrets Directive states that the holder of a trade secret cannot prevent others from reaching out the same knowledge and information by way of (a)“independent discovery or creation”, (b)“observation, study, disassembly or testing of a product or object that has been made available to the public or that is lawfully in the possession of the acquirer of the information who is free from any legally valid duty to limit the acquisition of the trade secret;”, or (c)“any other practice which, under the circumstances, is in conformity with honest commercial practices”. As observed by the wording of the sub-paragraph (b), reverse engineering is only permitted if the subject is “free from any legally valid duty to limit the acquisition of the trade secret”. Recital 16 clarifies this further by explicitly stating that “reverse engineering of a lawfully acquired product should be considered as a lawful means of acquiring information, except when otherwise contractually agreed. The freedom to enter into such contractual arrangements can, however, be limited by law”. In other words, trade secret holders can prohibit reverse engineering via contractual clauses, and this is what usually happens in practice except when this practice is prohibited by law.

2.3.2.1 “Public Interest” as an exception

The Trade Secrets Directive designates public interests and pursuit of other legitimate interests as exceptions of the trade secrets protection. Following Articles 1(2)(a) and 5(a), freedom of expression and information as well as the pluralism of media cannot be limited by trade secrets. Therefore, journalists are protected against legal claims based on trade secrets. Recital 19 reinforce Article 1(2)(a) by stating that not restricting the Article 11 of the Charter of Fundamental Rights of the European Union is Essential, “in particular with regard to investigative journalism and the protection of journalistic sources.” Moreover, Article 5(b) protects whistle-blowers since “revealing misconduct, wrongdoing or illegal activity” is exempted from the scope of trade secrets protection “provided that the respondent acted for the purpose of protecting the general public interest.” Lastly, trade secrets claims are also limited with the rights of employees. In this regard, Article 1(3) draws the boundaries of trade secrets protection in relation to mobility of employees whereas Articles 1(2)(d), 3(1)(c), and 5(c) address the activities of trade unions.

Most importantly, Article 1(2)(b) of the Directive allows “the application of Union or national rules requiring trade secret holders to disclose, for reasons of public interest, information, including trade secrets, to the public or to administrative or judicial authorities for the performance of the duties of those authorities.” Thereby, “public interest” gives way to algorithmic transparency as long as legal certainty is maintained. For instance, the EU Commission accessed to Google’s algorithms for an antitrust probe and found out that they were favouring Google Shopping[47] against its competitors. Here, it is worth repeating that

[47] See; EUROPEAN COMMISSION Competition CASE AT.39740 Google Search (Shopping) ANTITRUST PROCEDURE Council Regulation (EC) 1/2003. Available here:
https://ec.europa.eu/competition/antitrust/cases/dec_docs/39740/39740_14996_3.pdf (Accessed March 2021)

requiring public disclosure of the algorithms is possible as well. By the same token, Article 1(2)(c), backed-up by Recital 11, precludes the Directive from affecting "the application of Union or national rules requiring or allowing Union institutions and bodies or national public authorities to disclose information submitted by businesses which those institutions, bodies or authorities hold pursuant to, and in compliance with, the obligations and prerogatives set out in Union or national law". Furthermore, Recital 11 names the Regulation (EC) No 1049/2001 regarding public access to the documents of the Union institutions to exemplify such rules.

To conclude, "public interest" is a legitimate reason that enables the EU and the Member States to open the "legal black box" of the IP rights and require algorithmic transparency towards the public authorities as well as the public. For instance, "freedom of information laws" might be a common example for such a legal framework.

2.3.2.2 National Law as a legal basis for algorithmic transparency

The legal frameworks of the Member States can provide a basis for algorithmic transparency. For instance, source codes and other constituents of algorithmic systems that are procured from the private sector can be included in the information that is provided pursuant to freedom of information requests. However, Recital 18 of the Trade Secrets Directive states that the public authorities are not relieved from the "confidentiality obligations to which they are subject in respect of information passed on by trade secret holders, irrespective of whether those obligations are laid down in Union or national law." As named in the Recital, such confidentiality obligations can stem from, inter alia, the Directives 2014/23/EU about concession contracts, 2014/24/EU on public procurement, and 2014/25/EU about procurement by entities operating in the water, energy, transport, and postal services sectors. Therefore, sectorial regulations and different national practices can prevent the source codes and other relevant data from being disclosed.

As a matter of fact, there seems to be a confusion as well as a fragmentation in the practice regarding algorithmic transparency. For instance, the Digital Republic Act[48] in France constitutes a good example despite its shortcomings. Within this Act, new obligations were introduced in the Code of relations between the public and the administration (CRPA)[49]. Accordingly, the CRPA requires the publication of source codes of the algorithms used by public administrations, as well as transparency regarding the algorithmic decisions concerning individuals. These provisions were further supplemented by the "Decree n° 2017-330 of 14 March 2017 on the rights of individuals subject to individual decisions taken on the basis of algorithmic processing." According to Article L.300-2 of the CRPA, the source codes of the algorithms used by the administration constitutes "communicable documents."

[48] LOI n° 2016-1321 du 7 octobre 2016 pour une République numérique.

[49] Code Des Relations Entre Le Public et l'administration. Available here: https://www.legifrance.gouv.fr/codes/texte_lc/LEGITEXT000031366350?init=true&page=1&query=.+L.+312-1-1&searchField=ALL&tab_selection=all (Accessed Aprıl 2021)

Peculiarly, the “parcoursup case” in France exemplified how transparency could be tailored for a specific field given that a *lex specialis* applies. Parcoursup is a digital platform which was created by the Law n°2018-166 of 8 March 2018 to pre-register students in the higher education institutions. Although the basic algorithm of the parcoursup was published online[50], the “local algorithms” remained secret as the platform allows institutions to introduce their own selection criteria. Being an algorithmic system utilized in the context of admission decisions for higher education institutions, the Code of Education[51] governs the use of it. Following the Article L. 612-3 of the Code of Education, the right to obtain information regarding the criteria, procedures and pedagogical reasons applied to a final decision is solely reserved for the concerned applicants themselves. Following a claim filed by a student union to obtain the algorithm and the source code of the parcoursup system, the first instance Court ordered that these elements should be communicated to the student unions[52]. However, the *Conseil d’Etat* (the highest Court for the administrative matters) overruled this decision stating that the CRPA provision was not applicable in this case although noting that the “Decree of 19 March 2019” requires higher education institutions to release the general criteria used in their selection process[53].

On 15 January 2020, the *Conseil d’Etat* requested a preliminary ruling (*Question Prioritaire de Constitutionnalité)* [54] on the constitutionality of the L. 612-3 of the Code of Education, from the Constitutional Council (*Conseil Constitutionel)* considering that it limits the communication regarding the algorithmic decisions on the admission of students. In response, the Constitutional Council held[55] that the Article was in compliance with the Constitution, noting that higher education institutions are obliged to publish the criteria used for reviewing the applications once the national pre-registration procedure is completed and “the report must specify the extent to which algorithmic processing was used to carry out this examination and respect the privacy of applicants.” (FRA, 2020, p. 46) Thereby, this long judicial process revealed the algorithmic transparency modalities under the French legal framework concerning education as: a) publication of the basic source-code b) a right to explanation regarding a decision that can be exercised by the affected individual (candidate) and c) an ex-post general explanation towards the public, including the role of algorithmic systems in the decision process.

Another interesting case-law took place in Sweden, where the municipality of Trelleborg uses a process which they call “robotic automation” to handle the applications for

[50]The basic algorithm of the parcoursup platform is available here: https://www.enseignementsup-recherche.gouv.fr/cid130453/parcoursup-publication-du-code-informatique-des-algorithmes.html (Accessed April 2021)

[51]Code de l'éducation is available here : https://codes.droit.org/PDF/Code%20de%20l'%c3%a9ducation.pdf (Accessed April 2021)

[52] ‘Tribunal Administratif de La Guadeloupe N° 1801094’, 4 February 2019; V Thibault Douville, ‘Parcoursup à l’épreuve de La Transparence Des Algorithme’, 2019

[53] Le Conseil d’État, ‘Conseil d’État, 12 juin 2019, Université des Antilles,n°427916 & 427919’, 12 June 2019, Available here : https://www.conseil-etat.fr/ressources/decisions-contentieuses/dernieres-decisions-importantes/conseil-d-etat-12-juin-2019-universite-des-antilles (Accessed April 2021)

[54] Le Conseil d’État, ‘Base de jurisprudence, Decision n° 433296’.

[55] Conseil Constitutionnel, Décision 2020-834 QPC, 3 avril 2020.

financial aid. Following the freedom of information request of a journalist who want to see the source code of the software, the municipality claimed that it was a trade secret. On this dispute, the administrative Court decided that the code was not a trade secret but an official document which should be disclosed. Following the ruling, the code that municipality shared was made of "136,000 lines of rules, spread out across 127 XML files" including some files containing older, unused rulesets as well as the social security numbers and names of approximately 250 people (Algorithm Watch, 2020, p. 245). Notably, if taking appropriate measures for safeguarding such personal data was technically unfeasible, privacy and data protection laws would appear as a counterweight against transparency requirements. Yet, removal or anonymization of such data was a simple solution in this case which was neglected. Therefore, the moral of the story seems to be the need for clarifying which kind of information should be disclosed in which format.

Lastly, according to a report on the usage of algorithms by the public sector in Czechia, Georgia, Hungary, Poland, Serbia, and Slovakia, none of these countries met a good transparency standard as access to source codes or their algorithmic parts were rejected based on intellectual property rights or security reasons (E-Państwo Foundation, 2019).

As seen by these examples, there is not a uniform practice concerning algorithmic transparency in the European countries, neither in terms of which kind of data to share in freedom of information requests nor in terms of which modalities of transparency are deemed appropriate for which administrative sectors.

2.4 Chapter Conclusion: Tailoring Transparency for the Algorithmic Systems

Establishing transparency regarding algorithmic systems is a necessary policy goal that requires a regulatory intervention in a well-balanced and technically feasible manner. According to some scholars, although the problems are real, algorithmic transparency would not work due to "important computer science reasons" and solutions like algorithmic audits will only create the illusion of clarity although such a clarity is technically impossible (Desai & Kroll, 2017, p. 4). While dealing with technical opacity, "we should consider whether the algorithm presents such a great risk that limitations to its use are justified." (Buiten, 2019, p. 58) Indeed, these are legitimate concerns. Nonetheless, the technical challenges cannot be a valid reason to give up transparency. Instead, the limitations of the available technics should inform the lawmakers for drafting realistic legal frameworks and effective enforcement mechanisms. In this respect, opaque algorithmic systems that pose an acceptable amount of risk could be tolerated in proportion to their societal benefits just as cars are not prohibited because they cause accidents. Even in such cases, a form of transparency can be maintained by informing the end-users or persons affected regarding the qualities (including the opacity) of a given system with intelligible explanations and warnings.

On the other hand, opaque systems could be prohibited in certain fields or use cases where transparency is absolutely essential for safeguarding fundamental human rights, or crucial societal interests like the rule of law and the integrity of democratic processes. After all, a meaningful level of transparency should be mandatory for establishing accountability, safeguarding rights and liberties as well as mitigating the societal concerns, except for the limited cases where meaningful transparency is technically unrealistic, and the posed risks are insignificant.

For starters, it is necessary to shed light on the legal obscurity regarding the algorithmic systems starting from the IP rights and trade secrets in particular. Although the EU Trade Secrets Directive recognizes the disclosure of trade secrets to authorities or to the general public given that public interests are pursued, it still sets strong legal barriers in front of the algorithmic systems. According to the Directive, infringement actions against journalists or whistle-blowers who disclose the algorithms that a firm utilizes should not be successful in the European Courts as long as the disclosure serves the "public interest". Yet, it is worth noting that counting on the courage of journalists or whistle-blowers and a vague term such as "public interest" is still a slippery slope. Therefore, imposing some transparency on the private sector algorithms that could have serious societal or economic impacts might be a good idea besides establishing uniformity in the single market regarding the transparency of the public sector.

Considering the chaotic situation today, new European Regulations and legislative updates that are in tune with the technical challenges seem appropriate for achieving a meaningful level of algorithmic transparency. For this purpose, lawmakers should clearly define the required modalities of transparency in the sense that who should be given which information, when and in which format, starting from establishing effective enforcement agencies that would have access to the source codes and other relevant data of the algorithmic systems.

CHAPTER 3: ALGORTIHMIC TRANSPARENCY IN THE EU

The European discourse on regulating digital technologies is pretty ambitious in general. The EU claims to differ from the other global powers like the US and China by having a "human centric" approach that would safeguard rights and freedoms. For this purpose, the EU tries to develop regulatory policies which set out principles and rules for various forms of data processing via algorithmic systems. Arguably, this journey commenced with the GDPR which covers "automated data processing." This Regulation finally introduced "transparency principle" to the legal framework of the EU concerning algorithmic systems. Since then, "transparency" has been the focal principle in the regulatory enedeavours. Today, the EU aspires to continue its journey by establishing legal frameworks for the different uses of algorithmic systems and setting the standards, perhaps for the rest of the world as well. In this context, the DSA and the AI Act are especially important proposals that are designed for addressing specific categories of algorithmic systems. Therefore, this Chapter contains an analysis on how these proposed regulations might recalibrate algorithmic transparency in the EU.

3.1 Transparency in the GDPR

Although the GDPR is commonly known as a Regulation that concerns data collection and privacy, its scope actually covers all forms of "data processing" as it aims to regulate the whole life cycle of personal data. Article 4(2) defines "processing" in a broad way that covers "any operation or set of operations which is performed on personal data." Data processing via algorithmic systems fall within the material scope of the GDPR as Article 2(1) states that the Regulation applies to processing by "automated means."

It can be said that there are different domains in the GDPR from which some sort of an algorithmic transparency requirement can be derived. Firstly, Article 5 (a) lays down the fundamental principles that apply to all data processing activities which are: lawfulness, fairness, and transparency. Thereby, other principles in the Article 5 are interacting with the transparency principle. Namely, purpose limitation, data minimization, accuracy, storage limitation, integrity, confidentiality, and accountability are the other fundamental principles that the GDPR is built upon. Although the principles are interwoven in a way that they complement each other, it would not be wrong to say that transparency is the *primus inter pares* since it enables the application of the other principles. In fact, to assess whether a data processing activity is lawful, fair, accurate, proportional, and limited with the purpose or whether it establishes accountability, one certainly needs some sort of a transparency. On the other hand, being transparent about these qualities of a data processing activity is what defines and elaborates transparency principle. For instance, specifying the purpose and communicating it to the data subject is a form of transparency that is foreseen in the GDPR.

Following Recital 58, the transparency principle requires that "any information addressed to the public or to the data subject" to be "concise, easily accessible and easy to understand" in a way that "clear and plain language" and, where appropriate, "visualisation" is used. Thereby, the principle of transparency in the GDPR focuses on the "data subject" or the "public". Hence, intelligibility is underlined. As intelligibility requires that a provided information should be meaningful for the receiver of it, a layered approach to transparency is adopted. Following this approach, the children are further protected by requiring a clear and plain language that children could easily understand. Interestingly, the Recital 58 states that using electronic formats like websites is of particular relevance in situations where "the proliferation of actors and the technological complexity of practice makes it difficult for the data subject to know and understand whether, by whom and for what purpose personal data relating to him or her are being collected, such as in the case of online advertising." As the wording of the recital addresses "data collecting", one might argue that the transparency principle should be interpreted narrowly as if it only concerns informing individuals about the data collection. Therefore, it is worth repeating that the transparency principle is meant to apply to other data processing activities as well.

3.1.1 Transparency and the Rights of Data Subjects

Another domain of the GDPR that asserts some form of an algorithmic transparency is the "rights of data subjects" that are listed in the Articles from 12 to 23. In fact, some of these rights triggered "the right to explanation" debate concerning automated decision making.[56] In this respect, scholars Wachter, Mittelstadt and Floridi analyze what could be meant by an "explanation" in the context of an automated decision. As they categorize, an explanation may refer to the *system functionality*, as "the logic, significance, envisaged consequences, and general functionality" of an ADM, e.g. "the system's requirements specification, decision trees, pre-defined models, criteria, and classification structures" or to the *specific decisions*, as "the rationale, reasons, and individual circumstances of a specific automated decision", e.g., "the weighting of features, machine-defined case-specific decision rules, information about reference or profile groups". (Wachter, et al., 2017, p. 78) Indeed, this distinction is very crucial as it calibrates the modality of the required transparency regarding the algorithmic systems.

Among the rights of data subjects, there are several Articles that touch upon the algorithmic transparency issue. Firstly, the notification duties of data controllers under Articles 13 and 14, which are commented upon by Recitals 60-62 provide a pedestal for algorithmic transparency. Following Recital 60, to ensure fair and transparent processing, the controller should inform the data subject about the existence of the processing (including profiling), the purposes, the consequences (of profiling) and any further information necessary considering "the specific circumstances and context" in which the data processing takes place. Article 13(2)(f) requires data controllers to provide the information "at the time when personal data are obtained" regarding " the existence of automated decision-making, including profiling

[56] See Chapter 1.

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." Thereby, the notification duties in the GDPR require an ex-ante transparency regarding the system functionality.

Secondly, "the right to access" under Article 15 which is commented upon by Recital 63 might lead to a modality of algorithmic transparency. The Article provides a right to have "meaningful information about the logic involved" in the context of decision-making systems. In fact, Articles 13(2)(f), 14(2)(h) and Article 15(1)(h), are identical. They grant data subjects a right to be informed about "the existence of automated decision-making, including profiling" and "the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject". On the other hand, the right to access is different than the notification duty since it depends on the request of the data subject and does not have a time limitation. However, usage of the same language implies that an ex-post explanation of the system functionality instead of an explanation of the specific decision is sufficient. In other words, even if the data subject acts after a decision is taken, data controller would only be obliged to inform her with the general logic of the system and the "envisaged" consequences of the processing for her.

Last but not least, Article 22 of the GDPR addresses the automated decisions concerning individuals, including profiling. Accordingly, the data subject "shall have the right not to be subject to a decision based solely on automated processing, including profiling, which produces legal effects concerning him or her or similarly significantly affects him or her." Thereby, this Article goes beyond requiring transparency and prohibits automated decision making to some extent. However, the provision only applies if there is a "legal or similarly significant effect." According to the guidelines of the Working Party 29 (which is the ancestor of the European Data Protection Board), to affect someone "similarly significant", a decision must have a) the potential to significantly affect the circumstances, behaviour or choices, b) have a prolonged or permanent impact on the data subject or c) lead to exclusion or discrimination[57]. Moreover, as the wording of the article expresses that the right is limited to "…a decision based solely on automated processing…", a way to work around Article 22 might be inserting a "human in the loop". In this respect, the same guidelines require "meaningful human input" from a person with the necessary competence rather than a "token gesture" for the decisions to be not solely automated[58]. Following this logic, an algorithm like COMPAS[59] could be utilized in the EU since the decisions are not solely automated and the human inputs are made by someone with the relevant authority and competence such as a judge. Thereby, the GDPR seems to underestimate the problems that algorithmic systems might cause. In practice, decision support systems tend to take humans "under the loop" because the real person who "decides" does not necessarily know about the "reasoning" of the non-transparent

[57] See; ARTICLE 29 DATA PROTECTION WORKING PARTY, Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679 Adopted on 3 October 2017 As last Revised and Adopted on 6 February 2018.
[58] Ibid.
[59] Chapter 2.

algorithm and incline to trust it instead of risking a wrong decision by ignoring the systems input (Danaher, 2016). Considering this problem, measures like limiting the legitimate use cases of algorithmic systems and most importantly, designing a meaningful algorithmic transparency framework gains enormous importance for enabling humans to take independent decisions and therefore being accountable for their decisions.

In addition to the conditional prohibition of the automated decision making, Article 22(3) provides the data subjects with “at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision”. Accordingly, an intervention by a human who has the appropriate authority and who “undertakes a thorough assessment of the relevant data, including any additional information provided by the data subject” is a key element for establishing appropriate safeguards[60]. Following these due process rights, data subjects would need a reasoned decision in order to be able to contest it or even to ask for human intervention. Thus, this article can be interpreted as a right to an explanation of the specific decision concerning the data subject. Yet, although algorithmic transparency is recognized as an essential component for the due process, it is required in a very mild way. On this issue, the guidelines of the Working Party 29 states that the GDPR requires the controller to provide “meaningful information about the logic involved, not necessarily a complex explanation of the algorithm used or disclosure of the full algorithm[61].”

3.1.2 Inadequacies of the GDPR in terms of Algorithmic Transparency

The GDPR revolves around the data subjects and tries to provide them some information about the data processing activities. To be fair, it does introduce some modalities of algorithmic transparency. However, the scopes of the relevant provisions are pretty limited. Prima facie, what is made mandatory with the GDPR seems to be providing enough insight into the logic of an algorithmic system and the significance of that logic so that a data subject would have the context that is necessary to intelligently opt-in for an automated decision or to contest it. Disclosure of the algorithmic systems to the public or to individuals is not required under any circumstances since the IP rights overrule the data subjects’ rights. Therefore, transparency and due process rights regarding automated decisions are prone to be undermined in practice.

To make GDPR more effective, some scholars suggested a “right to reasonable inferences” that would require data controllers to proactively establish that an inference is reasonable. Accordingly, data controllers would be required to explain “(1) why certain data are normatively acceptable basis to draw inferences; (2) why these inferences are normatively acceptable and relevant for the chosen processing purpose or type of automated decision; and (3) whether the data and methods used to draw the inferences are accurate and statistically reliable.” (Mittelstadt & Wachter, 2019, p. 581) This way, the data subject would be able to practice its rights including the right to access or contest a subsequent decision based

[60] See; Supra note 57, p.27.
[61] Ibid, p. 25.

on the inference under Article 22 (3). As suggested by Wachter and Mittelstadt, this right could come into play for the "high risk inferences" that would be the ones drawn by big data analytics and which might damage the privacy or reputation of the data subject; or the ones being used for important decisions despite having low verifiability. Indeed, broadening the list of intelligible information that is communicated to data subjects might be a meaningful transparency modality that fosters fairness and accuracy.

Lastly, as admitted by the EU bodies, the enforcement of the GDPR did not live up to its full potential for now. One reason for this situation is the complicated enforcement model that gives the lead to the data protection authority of the State where a company has its establishment. As the European Commission underlines, "the largest big tech multinationals are established in Ireland and Luxembourg, the data protection authorities of these countries act as lead authorities in many important cross-border cases and may need larger resources than their population would otherwise suggest[62]." Moreover, the DPAs are tasked with ensuring fair, accountable, and transparent data processing without necessarily having a full access inside the legal and technical black boxes. After all, these seem to be the weak spots of the GDPR since effective enforcement is crucial for making sure that a regulation fulfils its promises.

3.2 Transparency in the Digital Services Act

The idea of having ex-ante regulations for the digital services came as a result of the emerging consensus on the inefficiency of soft-law approach and self-regulation. On top of the pressing policy goals like taking fake news under control for the governance of the pandemic and safeguarding the integrity of democratic elections, brutal terror attacks in Europe some of which were enabled by propagation of personal information and hate speech on social media caused the Member States to adapt laws that regulate social media and support an EU Regulation for this purpose. Hence, the Commission stepped in by proposing the Digital Services Act (DSA) which "lays down harmonised rules on the provision of intermediary services in the internal market." Indeed, the legal basis for the proposal is Article 114 of the Treaty on the Functioning of the European Union, which justifies the measures aiming to ensure the functioning of the European internal market.

As declared by the first Article, the aim of the DSA is to "contribute to the proper functioning of the internal market for intermediary services and to "set out uniform rules for a safe, predictable, and trusted online environment, where fundamental rights enshrined in the European Charter are effectively protected." For these purposes, the Regulation sets up a conditional exemption framework regarding the liability of intermediary service providers as well as rules on specific due diligence obligations tailored to certain specific categories of

[62]See; Brussels, 24.6.2020 COM (2020) 264 final COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT AND THE COUNCIL Data protection as a pillar of citizens' empowerment and the EU's approach to the digital transition -two years of application of the General Data Protection Regulation{SWD(2020)115 final, p 6.

service providers. With the adoption of the DSA, Articles from 12 to 15 of the E-Commerce Directive[63] will be deleted since the same issues are addressed within the new Regulation. The scope of the obligations that the DSA lays down includes all the actors that offer goods, information, or services in the EU, regardless of their place of establishment. Consequently, it applies to all digital service providers as a horizontal Regulation. On the other hand, the Commission took a two layered approach in the proposed draft by introducing additional obligations for the "very large online platforms" (VLOPs) that have a significant reach in the Union. Since the systemic risks posed by the online platforms are proportional to their sizes, the preference to monitor VLOPs more closely seems reasonable. In the current draft, the threshold for falling under the scope of these additional obligations is estimated to be having more than 45 million service recipients in the Union. Furthermore, this threshold is foreseen to be adjusted by the Commission in a way that it consistently corresponds to 10% of the Union's population. Not surprisingly, the objective criteria in the DSA mainly cover the American platforms which play a gatekeeping role in the digital economy like Google, Amazon, Facebook, Apple etc. In this respect, being a supporter of liberal market economy, the EU Commissioners never indicate the origin of these firms as the problem. While announcing the proposals, Thierry Breton, the European Commissioner for internal market and Margrethe Vestager, the executive vice president of the European Commission for a Europe Fit for Digital Age, kept repeating that "with bigger size, comes bigger responsibility".

3.2.1 The Enforcement Structure of the DSA

As for the enforcement, the initial draft of the DSA envisages the designation of "Digital Services Coordinators" (DSC) which will have a supervisory role besides a coordination task within the Member States as well as in the EU. Accordingly, all service providers are obliged to establish a single point of contact (Article 10) whereas providers which do not have an establishment in the Union but that offer services in the Union are obliged to designate, a legal or natural person as their legal representative in one of the Member States (Article 11) In case the service provider has a main establishment within the Union, the Member State in which the establishment resides will have jurisdiction to enforce the Regulation. If not, the Member State where a service provider appoints a legal representative will have jurisdiction, considering the function of legal representatives under this Regulation. In this regard, the Recital 76 iterates Article 40 by stating that "all Member States should, however, have jurisdiction in respect of providers that failed to designate a legal representative, provided that the principle of *ne bis in idem* is respected". As for the fines, Article 42 leaves it for the Member States to lay down the rules on penalties (not exceeding the 6% of the annual turnover of the service provider) applicable to breaches of the obligations.

[63] DIRECTIVE 2000/31/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 8 June 2000on certain legal aspects of information society services, in particular electronic commerce, in the Internal Market (Directive on electronic commerce)

Most importantly, the Regulation offers a stratified framework in terms of supervision, investigation, enforcement, and monitoring in a way that it is rather rigorous when it comes to the VLOPs. It provides for "enhanced supervision" in the event such platforms infringe the provisions of Chapter III of the Section 4 (Article 50) as well as the possibility for the Commission to intervene in case the infringements persist (Article 51). In such situations, the Commission can carry out investigations, including through requests for information (Article 52), interviews (Article 53) and on-site inspections (Article 54). It can adopt interim measures (Article 55) and make the commitments made by the VLOPs "binding" (Article 56), as well as monitor their compliance with the Regulation (Article 57). If very large online platforms are not complying with the decisions, the Commission can adopt non-compliance decisions (Article 58), issue fines up to the 6% of their annual turnover (Article 59) and impose periodic penalty payments (Article 60). After all, it seems like the Commission prefers a hands-on approach, drawing lessons from the rather ineffective enforcement structure of the GDPR.

3.2.2 The New Transparency Modalities under the DSA

The DSA proposal contains significant novelties for the algorithmic transparency in the EU. For starters, it requires transparency towards public authorities, independent auditors, and vetted researchers. However, whether the transparency provisions in the DSA enables the DSCs to access all the information regarding the algorithmic systems is a question that remains to be answered. Recital 64 of the DSA states that DSCs may require access to specific data including (but not limited to) "the data necessary to assess the risks and possible harms brought about by the platform's systems, data on the accuracy, functioning and testing of algorithmic systems for content moderation, recommender systems or advertising systems, or data on processes and outputs of content moderation or of internal complaint-handling systems within the meaning of this Regulation." However, the effectiveness of the DSCs could depend on the cooperation of the firms in practice. Worst case scenario, the Commission would be able to require access to the algorithms of the VLOPs in the context of on-site inspections (Article 54) or monitoring actions (Article 57) as remarked by the Recital 99. Yet, Article 31 (6) gives VLOPs a right to request the DSC of establishment or the Commission, to amend the request for access, "where it considers that it is unable to give access to the data requested because one of following two reasons: (a)it does not have access to the data; (b)giving access to the data will lead to significant vulnerabilities for the security of its service or the protection of confidential information, in particular trade secrets." Although the DSC or the Commission can insist on the request for access, the DSA foresees a negotiation that allows for a pushback by the online platforms especially in terms of not disclosing trade secrets. Therefore, even the public authorities could fail to penetrate the legal black box under the proposed DSA framework.

The most innovative transparency modality in the DSA is the requirement to provide access to data to vetted researchers "for the sole purpose of conducting research that contributes to the identification and understanding of systemic risks." Yet, following Article 31 (5), the Commission shall adopt delegated acts that lay down the specific conditions under which such sharing of data with vetted researchers can take place “taking into account the rights and interests of the very large online platforms and the recipients of the service concerned, including the protection of confidential information, in particular trade secrets, and maintaining the security of their service.” Therefore, trade secrets will be excluded from the information that is shared with the vetted researchers as well.

Eventually, the DSA foresees mandatory independent audits -at least once a year- for the algorithms that the VLOPs utilize (Article 28). According to the Recital 60, Auditors should “guarantee the confidentiality, security and integrity of the information, such as trade secrets, that they obtain when performing their tasks and have the necessary expertise in the area of risk management and technical competence to audit algorithms”. Thereby, this recital makes it clear that the auditors will have access to the trade secrets which could as well be the source codes of the algorithmic systems, training data of the ML models etc. Furthermore, Article 28 (4) obliges VLOPs to implement measures following the recommendations within the audit reports or justify not doing so. Accordingly, they should make their “implementing reports” public besides transmitting it to the regulatory authorities pursuant to Article 33 (2).

In brief, the DSA renders the algorithmic systems transparent for the independent auditors and maybe to the Commission (although with a huge possibility to pushback). Yet, the situation of the DSC’s remains blurry whereas vetted researchers seem to get a limited access which is curated by the platforms according to the initial draft.

3.2.3 Transparency Towards the Public

If the proposed draft is adopted as it is, there will also be new transparency obligations towards the public which corresponds to the more popular use of the word “transparency.” According to the draft, providers of intermediary services are obliged to document and publish detailed reports on any content moderation that they engage in following Article 13. These “transparency reports”, should include information relating to the removal and the disabling of information considered to be illegal content or contrary to the providers’ “terms and conditions”. In addition, VLOPs shall publish risk assessment reports, risk mitigation measures that are identified and implemented, audit reports and audit implementation reports. However, Article 33 (3) states that if the platform considers that the publication of these information “may result in the disclosure of confidential information of that platform or of the recipients of the service, may cause significant vulnerabilities for the security of its service, may undermine public security or may harm recipients” the platform is allowed to remove such information from the reports. In this case, the platform shall still transmit the complete reports to the “Digital Services Coordinator of establishment” and “the Commission”, accompanied by “a statement of the reasons for removing the information from the public reports.” Thereby, this article does not require platforms to disclose their “confidential information” like the source codes of algorithmic systems they might be using for content moderation to the public. Yet, they are required to be more transparent when it comes to the regulatory authorities.

3.2.4 Transparency Towards the Individuals

Finally, the DSA foresees some sort of "explanations" regarding certain types of algorithmic decisions. For instance, the online advertisement transparency requirements in the DSA are rather significant although their application in practice remains to be seen. At the minimum, Article 24 obliges all the online platforms to provide "(a) that the information displayed is an advertisement; (b) the natural or legal person on whose behalf the advertisement is displayed; (c) meaningful information about the main parameters used to determine the recipient to whom the advertisement is displayed." Remarkably, the requirement in the sub-paragraph (c) resembles the wording of the GDPR articles and recitals touching upon algorithmic transparency and it is open for an interpretation which could enable the "right to reasonable inferences" that researchers came up with in the context of the shortcomings of the GDPR. Thereby, the Article introduces a requirement to disclose the main parameters of the inferences that are drawn for targeting the individual for displaying certain advertisement. Notably, this Article could be strengthened by replacing "main" with "all" and adding the justification of the parameters and the data about the accuracy of such inferences to this list.

On top of the general transparency requirement concerning online advertisements, the proposed draft requires VLOPs to compile the advertisement they display and make it publicly available in a repository. As the Article 30(2) states, the available information should include: "d) whether the advertisement was intended to be displayed specifically to one or more particular groups of recipients of the service and if so, the main parameters used for that purpose; (e) the total number of recipients of the service reached and, where applicable, aggregate numbers for the group or groups of recipients to whom the advertisement was targeted specifically."   This way, the DSA demands explanations regarding the specific algorithmic decisions (including profiling) in the context of online advertisement and further enables journalists, lawyers, civil society organizations and public authorities to scrutinize the information provided by the VLOPs regarding the targeted advertisement.

Last but not least, Article 29 introduces a transparency requirement for the recommender systems of the very large online platforms including a right to choose or modify the parameters of such algorithmic systems. Accordingly, the platforms will be obliged to provide an option which is not based on profiling. However, the transparency modality required for the recommender systems seems to be limited with an ex-ante explanation about the system functionality within the "terms and conditions." Notably, this rule could be circumvented in practice if the platforms keep profiling people without tailoring their newsfeed for them or nudge them to opt-in for profiling by showing them irritant or totally irrelevant content. Therefore, these rules could be meaningless unless they are enforced effectively.

3.3 Transparency in the Proposed Artificial Intelligence Act

The proposed Artificial Intelligence Act (AI Act)[64] lays down a pretty detailed framework for certain algorithmic systems that are called as "artificial intelligence" (AI). In fact, defining AI is a challenging task since it is a dynamic term. In colloquial language, intelligence is often seen as whatever machines or animals have not done yet[65]. This creates an "AI effect" since people tend to move the bar for calling something "AI" higher as technology progresses. Therefore, it has been heavily debated whether having a general regulation on AI would be the right approach in the first place. For instance, the German AI Association drafted a Position Paper stating that regulating "AI" is not feasible since "a clear definition of the term that allows us to differentiate AI from already existing algorithms is missing." (KI Bundesverband, 2021, p. 11) Ultimately, the Commission chose to list the techniques and approaches that are defined as "AI" for the purposes of its AI Act, in order to overcome this dilemma.

Article 3(1) of the AI Act defines an 'artificial intelligence system' (AI system) as:

"software that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."

As for the initial draft, Annex I covers:

"(a) Machine learning approaches, including supervised, unsupervised and reinforcement learning, using a wide variety of methods including deep learning; (b) Logic- and knowledge-based approaches, including knowledge representation, inductive (logic) programming, knowledge bases, inference and deductive engines, (symbolic) reasoning and expert systems; (c) Statistical approaches, Bayesian estimation, search and optimization methods."

Since new techniques and approaches can emerge, Article 4 empowers the Commission to adopt delegated acts to amend the list of techniques and approaches listed in Annex I. Thereby, the AI Act is designed as a "living document." After all, rather complex algorithmic systems that are created with the listed approaches or techniques will be covered under the regulation.

Notably, Article 2 of the Regulation leaves the AI systems that are developed or used exclusively for military purposes; that are used by public authorities in a third country or an international organization in the framework of an international agreement for law enforcement and judicial cooperation; and the high-risk AI systems that are safety components of products or systems, or which are themselves products and regulated under the listed *lex specialis*, out

[64] COM (2021) 206 final 2021/0106 (COD) Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts {SEC(2021) 167 final} - {SWD(2021) 84 final} - {SWD(2021) 85 final} Brussels, 21.4.2021.

[65] See; Larry Tesler, CV: Adages & Coinages. Available here: http://www.nomodes.com/Larry_Tesler_Consulting/Adages_and_Coinages.html (Accessed April 2021)

of the scope of the AI Act[66]. Consequently, many serious use cases of AI systems that can pose a threat to fundamental human rights are not covered by the Act. Moreover, the majority of AI systems available today are exempted from this Regulation as well since they pose minimal or no risk to human rights or safety. Therefore, the AI Act seems to have a rather limited scope.

For the AI systems that are covered, the AI Act contains four distinct regulatory regimes which are tailored in proportion to the level of risk posed by the given AI system. These are; "unaccaptable risk" that is prohibited; "High-risk[67]" that is strictly regulated; "limited risk" that is subjected to transparency obligations, and "minimal risk" or "no risk" that are allowed a free rein. In other words, the Regulation treats the AI systems depending on the risk category that they belong to. For starters, Article 5 bans a small number of potential use cases of AI systems that are considered as a threat to human rights and the safety[68]. Consequentially, all the AI systems are expected to be transparent at least about their intended purposes and general functioning so that an assessment can be made regarding whether an AI system falls under this prohibition rule. Secondly, transparency is deemed as strictly necessary for mitigating the risks posed by the "high-risk" AI systems to fundamental rights and safety. For such systems, transparency in the sense of informing end-users and persons affected is foreseen as the solution along with the high-quality data, documentation, traceability, human oversight, accuracy, and robustness requirements. Lastly, the AI Act imposes limited transparency obligation for the other non-high-risk AI systems. Therefore, transparency has a very significant role in this framework and it has different modalities depending on the risk category, technological complexity as well as the subject who gets access to the information .

As for the enforcement, the Member States will designate national competent authorities being a "national supervisory authority", a "market surveillance authority" and a "notifying authority" for implementing the rules. Then, the notifying authorities will notify the "notified bodies." Moreover, the national supervisory authorities will represent the States in the "European Artificial Intelligence Board" that would be established under the AI Act. For the purposes of this thesis, instead of delving into the enforcement structure envisaged in the AI Act, it would suffice to note that it is rather complicated, and this might turn into a weak spot.

3.3.1 Transparency in the Broad Sense

In general, European Regulations contain transparency requirements towards the general public, end-users or persons affected. In the European Law parlance, the word

[66] As for the "high-risk AI systems" that are safety components of products or systems, or which are themselves products or systems, falling within the scope of the listed acts, Article 84 of the AI Regulation still applies.

[67] As for the designation of high-risk AI systems, Article 6 (2) refers us to the Annex III which can be updated by the Commission following Article 7. Remarkably, the list in the Annex III has a sectorial aspect besides a technological one. For instance, educational or vocational training is listed as a high-risk context. On the other hand, all remote biometric identification systems are considered high risk as well. Putting such AI systems on the market will be subject to strict requirements.

[68] An analysis of these cases falls out of the scope of this thesis.

"transparency" is mostly used to indicate this modality of transparency which could be called the "narrow meaning of transparency". On the other hand, "transparency" in a more general or broad meaning covers transparency towards third parties (public authorities, audit firms, academicians, certification bodies etc.) and other measures which enables transparency like documentation etc. In this sense, the Regulation contains many new forms of transparency. One of such novelties is the ex-ante conformity assessments that are required for high-risk AI systems. In this context, independent third parties (notified bodies) will be involved in the conformity assessment procedures of the remote biometric identification systems and the high-risk AI systems other than those related to products. Although only for a very limited category of AI systems, the AI Act foresees full transparency towards such independent third parties as well as the public authorities. Peculiarly, Article 70 draws the limits of such transparency as compliance with relevant legislation in the field, including the EU Trade Secrets Directive and states that "when public authorities and notified bodies need to be given access to ***confidential information or source code*** to examine compliance with substantial obligations, they are placed under binding confidentiality obligations (emphasis added)." Following Article 64, the market surveillance authorities "shall be granted ***full access*** to the training, validation and testing datasets used by the provider, including through application programming interfaces ('API') or other appropriate technical means and tools enabling remote access" and "where necessary to assess the conformity of the high-risk AI system (…) and ***upon a reasoned request***, the market surveillance authorities shall be granted ***access to the source code*** of the AI system (emphasis added)." Accordingly, the public authorities will be given access to the source codes and other relevant secretive information as being bounded by confidentiality obligations. However, since Article 23 concerning the cooperation with competent authorities addresses "the providers of high-risk AI systems", to what extent the providers of non-high-risk AI systems should be transparent towards the public authorities is rather doubtful. In conclusion, although the *de facto* situation may manifest itself differently due to ineffective enforcement or pushbacks from private firms, *de jure* situation would be clarified with the AI Act as relevant articles indicate full disclosure and require a transparent information flow towards the public authorities. Thereby, a significant modality of algorithmic transparency could be established with the adoption of this provision of the AI Act as it is. Yet, it is worth noting that the application of these rules will be problematic given the fact that the Regulation leaves the initial designation of the risk category to the provider of the AI system.

Apparently, the public authorities and the conformity assessment bodies will be able to open the "black box" of the high-risk AI systems. However, the legal obstacles will still overrule when it comes to the narrow meaning of transparency. As mentioned in the foreword of the AI Act, "the increased transparency obligations" will not disproportionately affect the right to protection of intellectual property that is enshrined in the Article 17(2) of the European Charter, since "they will be limited only to the minimum necessary information for individuals to exercise their right to an effective remedy and to the necessary transparency towards supervision and enforcement authorities, in line with their mandates." Therefore, transparency in the narrow meaning will remain quite restricted. Even so, the proposed Regulation foresees some transparency modalities towards the individuals as well.

3.3.2 Transparency in the Narrow Sense

The focality of transparency for the AI Act can be observed right from the first Article which concerns the subject matter of the Regulation. As Article 1 (c) says, the Regulation lays down "harmonised transparency rules for AI systems intended to interact with natural persons, emotion recognition systems and biometric categorisation systems, and AI systems used to generate or manipulate image, audio or video content." Indeed, a certain kind of transparency is required under title IV for the AI systems that pose a "limited risk." Following Article 52[69] (1), providers of AI systems shall ensure that "AI systems intended to interact with natural persons are designed and developed in such a way that natural persons are informed that they are interacting with an AI system, unless this is obvious from the circumstances and the context of use." For instance, users should be aware if they are interacting with a chat-bot so that they can take an informed decision on whether to continue or step back. Similarly, paragraph 2 obliges emotion recognition or biometric categorization systems to inform "of the operation of the system the natural persons exposed thereto". Unfortunately, the wording of this paragraph is rather unclear since such systems can differ drastically and there are not any criteria given for assessing how detailed the provided information should be. Finally, the third paragraph requires that users of AI systems that generate or manipulate content which would falsely appear truthful or authentic (such as "deep fakes") to disclose that the content was generated or manipulated.

As for the high-risk AI systems, the Regulation drafts a more detailed modality of transparency. Following Article 13 (1), such systems should be designed and developed in a way that ensures their transparent operation, which should suffice for the users to interpret the systems output and use it appropriately. The "appropriate type and degree of transparency" is assessed by achieving compliance with other relevant obligations set out in Chapter 3 of the same title which in fact includes measures which enable transparency in the broad sense like putting a "quality management system" in place, drawing up technical documentation, conducting conformity assessments, and keeping automatically generated logs. After all, it would not be erroneous to name this approach as "transparency by design." Indeed, Article 12 requires that high-risk AI systems to be designed and developed as capable of enabling the automatic recording of events (logs) whereas Article 14 requires them to include an appropriate human-machine interface tool for establishing human oversight. Moreover, the oversight requirement in the Regulation is pretty ambitious. The Article 14 (4) writes that an assigned individual shall be enabled to "fully understand the capacities and limitations of the high-risk AI system and be able to duly monitor its operation" and "remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system." Furthermore, the individual should be able "to intervene on the operation of the high-risk AI system or interrupt the system through a "stop" button or a similar procedure." Indeed, having a "kill switch" in case AI goes wrong is the most intuitive and necessary solution so that the person in charge does not have to trigger a short-circuit by spilling a glass of water as

[69] The Article 52 does not apply if the practice is authorized by law in the context of fight against crime.

a last resort. Ultimately, it is worth noting that these are not only artificial intelligence design and development rules but also human intelligence training requirements. For instance, how can a person "remain aware" about the automation bias and act accordingly needs to be elaborated. In any case, it is noteworthy that the AI Act complements the Article 22 of the GDPR by aiming to limit the delegation of intellectual tasks to automated systems, although both Regulations have very limited scopes.

After designing and developing the high-risk AI system in a way that enables transparency and documenting necessary information, comes the duty to inform end-users or persons affected. Article 13 (2) requires high-risk AI systems to be accompanied by "instructions for use" -in a digital format or otherwise- that include "concise, complete, correct and clear information that is relevant, accessible and comprehensible to users." Within this obligation, the calibre of algorithmic transparency is shaped by Article 13 (3) that lists the items which shall be specified in such instructions. Interestingly, some items in this long list aims to provide an overall insight to the innerworkings of the algorithmic systems. For instance, the instructions should specify "the characteristics, capabilities, and limitations of performance of the high-risk AI system (…); its intended purpose; the level of accuracy, robustness, and cybersecurity;(…); ***when appropriate***, specifications for the input data, or any other relevant information in terms of the training, validation and testing data sets used, taking into account the intended purpose of the AI system; (…) the human oversight measures… (emphasis added)". Besides such information that could provide a limited insight regarding the algorithmic system in question, the list also includes **"**its performance as regards the persons or groups of persons on which the system is intended to be used" which might be a significant information in the context of algorithmic bias.

Cardinally, for informing the public further and facilitating the work of the Commission and the Member States, providers of high-risk AI systems will be required to register their AI systems in a database that will be established and managed by the European Commission. Following Article 60 (3), the EU database will be open to the public access. As listed under the Annex VIII, the information in the database will include the electronic "instructions for use" of the AI systems as well.

Finally, the transparency policy of the proposed Regulation deals with the technical opacity problem in a rather ambiguous way. In this respect, Recital 47 clarifies that a certain degree of transparency should be required for high-risk AI systems that are incomprehensible or "too complex for natural persons" so that the users can "interpret the system output and use it appropriately." For this purpose, they should be "accompanied by relevant documentation and instructions of use and include concise and clear information, including in relation to possible risks to fundamental rights and discrimination, where appropriate." Thereby, deploying a "high-risk" black box algorithm which might pose risks to fundamental rights seems to be allowed as long as the instructions and documents are provided. Moreover, since the Commission exemplified "minimal risk" AI systems as "AI-enabled video games or spam filters", such algorithmic systems seem to be accepted as they are, regardless of their opaqueness. However, assuming that AI-enabled video games cannot be manipulative or harmful seems simply wrong. As the initial assessment about the risk categories is up to the developers of the AI systems, this example is rather misguiding.

3.4 Chapter conclusion

Transparency remains to be a central element of the European regulatory policies. However, the regulations that are in force and that are proposed seem to have significant shortcomings in terms of delivering this promise. Firstly, a “full transparency” requirement as disclosure of source codes is nowhere to be found in the EU regulations. Instead, limited algorithmic transparency modalities are preferred. This is a policy decision that might undermine the success of the regulations in terms of safeguarding human rights, democracy, and rule of law. On the other hand, full transparency is only one of the many modalities of algorithmic transparency and it would be wrong to give it more weight than it deserves. After all, the regulations do contain other meaningful modalities of transparency despite their shortcomings.

As for the GDPR, the prominent shortcomings can be listed as the limited scope of the Article 22, absence of a “right to reasonable inferences” and most importantly, effective enforcement. Actually, it can be said that the enforcement envisaged under the GDPR was a “half-measure” which sacrificed harmonized and strong enforcement for giving the firms the chance to pick up a safe haven. Thereby, the GDPR neither enabled an EU level enforcement nor allowed national authorities to act without waiting for the lead of the DPA of establishment.

The proposed regulations seem to complement the GDPR to an extent. In respect of the DSA, there is a legitimate two layered approach that introduces stricter obligations for the platforms that scale up the negative impacts due to their sizes. Moreover, the Commission seems to have a more hands-on approach in terms of enforcement. Most importantly, it introduces third party algorithmic audits which was long awaited. However, it is worth noting that this modality of transparency can easily turn into a box ticking exercise and may create an unearned trust.

Lastly, the AI Act seems to complement the European framework especially with the rules that require “transparency by design.” However, it has serious weaknesses as well, starting with its very narrow scope that excludes most of the use cases of complex algorithmic systems. Even when an algorithmic system falls under the scope of the proposed Regulation, most of the obligations concern the “high-risk” AI systems. In fact, the Regulation mainly relies on self-assessment and methods like certification which could end up as box ticking practices just like the third-party audits under the DSA. Therefore, it remains to be seen whether an effective enforcement could be achieved.

CHAPTER 4: POLICY RECCOMMENDATIONS

4.1 Conclusion

Using transparency as a regulatory tool is an old idea that can have many different forms. For one, Louis Brandeis spearheaded a certain kind of transparency movement in the US by calling for transparency requirements concerning the financial sector back in 1914. Since then, the principle of requiring extended transparency for specific sectors like e.g., finance, industry or nutrition have been a popular approach all over the world. Today, there exists a strong transparency demand targeting online services and "artificial intelligence" sector which are based on deploying and developing algorithmic systems. In fact, algorithmic transparency is especially important for safeguarding human rights, democracy, and the rule of law, considering that the algorithmic systems are "technical regulations" which contain serious policy decisions. Although transparency is not the only focal aspect of the debates revolving around regulating these sectors, it is an indispensable one. Scholars Fung, Graham, and Weil describe "targeted transparency" requirements as "a distinctive category of public policies that, at their most basic level, mandate disclosure by corporations or other actors of standardized, comparable, and disaggregated information regarding specific products or practices to a broad audience in order to achieve a public policy purpose." (pp. 37-38) and differ targeted transparency policies from "warnings" and "right to know policies" as follows: "…warnings provide information that is simple and prescriptive, targeted transparency provides information that is complex and factual. Whereas warnings urge users to take a particular course of action, targeted transparency encourages users to make reasoned judgments of their own. And whereas right-to-know policies aim to generally inform public discourse, targeted transparency aims to influence specific choices." (Fung, et al., 2007, p. 39) In other words, targeted transparency is merely a disclosure policy which counts on the individuals to judge the disclosed information and not a monkey wrench solution for every given problem. While dealing with algorithmic systems, it might be useful to go beyond this paradigm and conceptualize a transparency policy in a broader sense. For starters, it is possible to define six different modalities of transparency which can correspond to different policy options in combination or as themselves. Accordingly, the following approaches can be adopted:

1) **Laissez-faire:** It is possible to trust the entities that develop or deploy algorithmic systems blindly and rely on self-regulation. In this case, algorithmic systems can be veiled by intellectual property rights, trade secrecy or other reasons. In other words, opacity would be accepted.

2) **Transparency towards the individuals:** It is possible to require certain algorithmic systems to provide information and intelligible explanations to the end-users and persons affected. In addition, individuals can exercise their right to information in a proactive manner with individual information requests depending on the legal framework. In fact, the word "transparency" is mostly used for referring to these two modalities in the EU parlance. Accordingly, whether an *ex-ante* or an *ex-post*

explanation is given and whether an explanation of a specific decision or a general information about the functionality of a system can differ. The EU already has a legal framework that provides intelligible explanations without disclosing the source codes and relevant data (the GDPR, P2B Regulation, consumer protection law etc.) which will be pushed further with the proposed Regulations. Under these regulations, the individuals are provided some intelligible insights to the system, instead of full transparency. As for the freedom of information laws, there seems to be a fragmentation within the Union. As for today, these transparency modalities do not always require disclosing source codes or other relevant data to the individuals. Often times, they remain veiled by intellectual property rights, trade secrecy or other reasons, especially when it comes to the private sector.

3) **Transparency towards the public authorities:** It is possible to explicitly allow public authorities to access the source codes and other relevant data of algorithmic systems for novel purposes. When this is done as an *ex-post* measure, entities can push back against the requests for access to some extent. Furthermore, even if the access is legally justified, it is not easy to apply this in practice especially when the entity in question is not within the jurisdiction of a given authority. In fact, the EU wants to establish such an information flow with the new Regulations, noting that the public authorities will remain under confidentiality obligations. Moreover, this modality of transparency can be introduced as an *ex-ante* measure by requiring pre-market authorization for some algorithmic systems.

4) **Transparency towards trusted third parties:** It is possible to require entities that develop or deploy algorithmic systems to grant access to trusted third parties like audit firms, certification bodies, vetted academicians etc. Indeed, these are foreseen in the proposed EU Regulations in a rather limited way.

5) **Targeted full transparency:** It is possible to require full algorithmic transparency for certain sectors in the sense that source codes and other relevant data regarding the algorithmic systems will be publicly accessible. Such policies often target the public sector like the Digital Republic Act in France. This policy can be adopted on a voluntary basis as well. For instance, municipalities like Barcelona[70] and Amsterdam[71] are making the source codes and relevant data of the algorithmic systems they use available to public.

6) **Absolute transparency:** It is possible to leave algorithmic systems out of the scope of intellectual property rights and trade secrets as it was initially discussed by the supporters of free-software and open-source movements. In this scenario, source codes of all the algorithmic systems would be open for everyone to examine (and possibly to

[70] The GitHub algorithmic repository of the Barcelona municipality is available here: https://github.com/AjuntamentdeBarcelona (Accessed April 2021)

[71] Algorithm Register of the Amsterdam Municipality is available here: https://algoritmeregister.amsterdam.nl/en/ai-register/ (Accessed April 2021)

reproduce and develop). On the other hand, a limited absolute transparency policy can be adopted on a voluntary basis as well. For instance, Barcelona Municipality is a part of the "public money-public code" campaign of the free software foundation and uses free-software.

In this respect, Lessig writes that "naked transparency" is a bad idea since people are not necessarily informed when huge amounts of data are disclosed, but they are often times careless and open to believing in wrong narratives because the public is "too smart to waste its time focusing on matters that are not important for it to understand" adding that "the ignorance here is rational, not pathological." (Lessig, 2009) For instance, this is the case with the "terms and conditions" or "privacy policies" of online services that people "agree" or "consent" to nowadays, which would require, on average, 25 days a year to read[72]. Therefore, adopting an "absolute transparency" policy by itself would not suffice to establish meaningful transparency in general as it does not always provide intelligibility for the general public. On the other hand, it should be remembered that it is one of the strongest tools in this arsenal since making the relevant information available enables scrutiny and openness as coders, journalists, lawyers, researchers, policy makers and civil society organizations can work on it.

4.2 Policy Recommendations

In light of the analysis above, the following recommendations could be useful for achieving the ambitious policy goals of the EU and establishing a meaningful level of transparency regarding algorithmic systems:

- ***Develop a framework for specifying the format and the elements of the information that should be shared following a "freedom of information request" to make sure that trade secrecy or other intellectual property rights cannot hinder the "right to information."***

  Notably, regardless of whether the source codes and the relevant data are fully disclosed, simplified, and legible explanations should be provided for the individuals seeking information (as well as the end users and persons affected). As no society only consists of experts in information technologies, there would always be a need for easier means of interpretation like natural language explanations or visualisations. Besides specifying the required qualities of the shared information, this framework should also indicate how to safeguard the right to data protection while sharing such relevant data.

- ***Establish public repositories for algorithmic systems that are utilized by the public sector and make public disclosure of source codes that are not kept secret for***

[72] See, European Data Protection Supervisor, Opinion 8/2016 on coherent enforcement of fundamental rights in the age of Big Data (2016), p.13. Available here: https://edps.europa.eu/sites/edp/files/publication/16-09-23_bigdata_opinion_en.pdf (accessed April 2021)

legitimate reasons e.g., national security or concerns of "gaming the system" mandatory.

In addition to this, requiring the use of free software could be the optimal solution for fostering interoperability, innovation, as well as competition among suppliers. This way, anyone with the necessary skills could check whether the computer implementation of an algorithm is correct, examine if there are negative societal impacts like unfair discrimination, or even improve the systems. As for the algorithmic systems based on ML, access to documents regarding the training process and the underlying data without infringing data protection or privacy rights might help with assessing the qualities of the system. Furthermore, specific information regarding the intended purpose of the algorithmic system as well as validation and follow up processes could be required to be made public. For the dynamic models that change according to new patterns in the data, periodical public disclosures can be required. This way, some hidden policy trade-offs of deploying opaque algorithms can be exposed to the public for democratic scrutiny.

- ***Consider introducing a requirement to disclose the source codes and other relevant data for selected sectors besides the public sector unless there are serious security concerns.***

Potential implications of a "full transparency" requirement in the sense of publicly disclosing the source codes and other relevant data of algorithmic systems utilized in certain sectors (e.g., private transportation applications) where disclosure could serve the "public interest" should be assessed. Thereby, policy makers should start considering targeted full transparency as a regulatory tool for the sectors that are increasingly becoming "digital public infrastructures." Besides a requirement for certain sectors, public disclosure could also be a sanction for not complying with the rules as well. This would foster fair competition and innovation besides the democratic oversight.

- ***Update the Regulation (Eu) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).***

Amending the GDPR and introducing new rights for the "data subjects" could be necessary for establishing a more transparent data processing ecosystem. For starters, the enforcement mechanism which is rather a "half-measure" could change. In fact, the existing structure neither enables an EU level enforcement nor allows national authorities to act without waiting for the DPA of establishment. Furthermore, introducing a general "right to reasonable inferences" can strengthen the algorithmic transparency framework of the EU in an overarching way.

- ***Strengthen the Proposal for a Regulation of the European Parliament and Of the Council on a Single Market for Digital Services (Digital Services Act) and amending Directive 2000/31/EC, COM (2020) 825 final 2020/0361 (COD) of 15.12.2020.***

  The current draft of the DSA renders the algorithmic systems transparent for the independent auditors and maybe to the Commission (although with a huge possibility to pushback). Yet, the situation of the DSC's remains blurry whereas vetted researchers seem to get a limited access which is curated by the platforms. Regardless of whether the targeted transparency and public disclosure policies are preferred or not, granting public authorities, academicians, and independent auditing firms an extended access to the algorithmic systems and other relevant data is indispensable. This way, experts that are tasked with analysing the algorithmic systems would ensure that there is at least some scrutiny that makes transparency "meaningful". The expert audits should cover the impacts of algorithms on human rights, democracy, and the rule of law in detail. Adopting the DSA by making sure that the wordings of the clauses concerning transparency towards the public authorities and enforcement would not render it toothless could achieve establishing this modality of algorithmic transparency.

- ***Strengthen the Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts, Com (2021) 206 Final Of 21.4.2021.***

  Reinforcing the proposed AI Act would be better for establishing a robust framework. For starters, an effective and less complex enforcement model could be established, drawing lessons from the shortcomings of the GDPR enforcement model. For having a "one-stop-shop" for the AI developers and deployers as well as avoiding fragmentation, an EU Level enforcement mechanism might be the best solution. Having an EU-Level regulatory body would also enable a strict and fast pre-market authorization policy. Thereby, the ex-ante self-assessment and certification policies proposed in the AI Act could be supported by requiring certain high-risk AI applications to be tested and authorized within this new EU Body.

Bibliography

Algorithm Watch, 2020. *Automating Society,* Berlin: AlgorithmWatch gGmbH.

Ananny, M. & Crawford, K., 2018. Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *new media & society,* pp. 973-989.

Assange, J., 2012. *CYPHERPUNKS: Freedom and the future of the internet.* New York and London: OR Books.

Barlow, J. P., 1996. *A Declaration of the Independence of Cyberspace.* [Online]
Available at: https://www.eff.org/cyberspace-independence

Borgesius, F. Z., 2018 . *Discrimination, artificial intelligence, and algorithmic decision-making,* Strasbourg: Published by the Directorate General of Democracy Council of Europe .

Brandeis, L., 1914. *Other People's Money (Ch. 5: What Publicity Can Do).* s.l.:https://louisville.edu/law/library/special-collections/the-louis-d.-brandeis-collection/other-peoples-money-chapter-v.

Brauneis, R. & Goodman, E. P., 2018. Algorithmic Transparency for the Smart City. *THE YALE JOURNAL OF LAW & TECHNOLOGY,* 103 (20), pp. 103-176.

Broussard, M., 2018. *Artificial Unintelligence: How computers misunderstand the world.* s.l.:MIT Press.

Brownsword, R., 2005. Code, control, and choice: why East is East and West is West.. *Legal Studies,* 25(1), p. p. 1–22..

Buiten, M. C., 2019. Towards Intelligent Regulation of Artificial Intelligence. *European Journal of Risk Regulation,* 10(1), pp. 41-59..

Burrell, J., 2016. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society,* January-June, pp. 1-12.

Cambridge Dictionary, 2021. *Cambridge English Dictionary.* [Online]
Available at: https://dictionary.cambridge.org/dictionary/english/transparency

Cardon, D. & Crépel, M., 2019. *Algorithms and Territorial Regulation,* s.l.: books&ideas.

Castelluccia, C. & Métayer, D. L., 2019. *Understanding algorithmic decision-making: Opportunities and challenges,* Brussels: European Parliamentary Research ServiceScientific Foresight Unit (STOA).

Cleg, N., 2021. *You and the Algorithm: It Takes Two to Tango.* [Online]
Available at: https://nickclegg.medium.com/
[Accessed April 2021].

Crawford & Kate, M. A. a., 2016. Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability.. *New Media & Society,* p. 973–989.

Danaher, J., 2016. The Threat of Algocracy: Reality, Resistance and Accommodation, 29245. *PHILOSOPHY & TECHNOLOGY ,* 29(245).

DARPA, 2016. *Broad Agency Announcement: Explainable Artificial Intelligence (XAI)DARPA-BAA-16-53,* Arlington: DARPA Information Innovation Office.

DARPA, 2018. *DARPA Announces $2 Billion Campaign to Develop Next Wave of AI Technologies.* [Online]
Available at: https://www.darpa.mil/news-events/2018-09-07
[Accessed March 2021].

Daston, L., 2013. *How Reason Became Rationality.* [Online]
Available at: https://www.mpiwg-berlin.mpg.de/research/projects/DeptII_Daston_Reason
[Accessed April 2021].

Datatilsynet, 2018. *Artificial intelligence and privacy,* s.l.: Datatilsynet (The Norwegian Data Protection Authority).

Desai, D. R. & Kroll, J. A., 2017. TRUST BUT VERIFY: A GUIDE TO ALGORITHMS AND THE LAW. *Harvard Journal of Law & Technology,* 31(1).

Diakopoulos, N., 2014. “Algorithmic Accountability.”. *Digital Journalism, vol. 3, no. 3,,* p. pp. 398–415.

Diakopoulos, N., 2016. *We need to know the algorithms the government uses to make important decisions about us.* [Online]
Available at: https://theconversation.com/we-need-to-know-the-algorithms-the-government-uses-to-make-important-decisions-about-us-57869
[Accessed April 2021].

Diega, G. N. L., 2018. Against the Dehumanisation of Decision-Making: Algorithmic Decisions at the Crossroads of Intellectual Property, Data Protection, and Freedom of Information. *JIPITEC,* 3(1).

Dumas v. Liebert (1868).

Edwards, L. & Veale, M., 2017. Slave to the algorithm? Why a ‘right to an explanation’ is probably not the remedy you are looking for. *DUKE LAW & TECHNOLOGY REVIEW,* Vol. 16(No. 1).

E-Państwo Foundation, 2019. *alGOVrithms,* s.l.: s.n.

Fink, K., 2018. Opening the government’s black boxes: freedom of information and algorithmic accountability. *Information, Communication & Society,* 21(10), pp. 1453-1471.

FRA, 2015. *Freedom to conduct a business: exploring the dimensions of a fundamental right,* Luxembourg: Publications Office of the European Union.

FRA, 2020. *Artificial Intelligence, Big Data and Fundamental RightsCountry ResearchFrance2020Report provided to FRA under contract D-SE-19-T02,* s.l.: European Union Agency for Fundaamental Rights.

Fung, A., Graham, M. & Weil, D., 2007. *Full Disclosure: The Perils and Promise of Transparency.* New York: Cambridge University Press.

G’sell, F., 2019. *Personalization of the Law : a French Perspective.* Chicago, Legal Challenges of the Data Economy conference.

Goodman, B. & Flaxman, . S., 2016. *European Union regulations on algorithmic decision-making and a “right.* [Online]
Available at: https://arxiv.org/pdf/1606.08813.pdf
[Accessed March 2021].

Gunning, D. & Aha, D. W., 2019. DARPA’s ExplainableArtificial Intelligence Program. *AI MAGAZINE,* Summer(Deep Learning and Security), pp. 44-58.

Han, B.-C., 2015. *The Transparency Society.* Palo Alto, United States: Stanford University Press.

Hildebrandt, M., 2013. *“Slaves to Big Data: or Are We?”.* [Online]
Available at: https://works.bepress.com/mireille_hildebrandt/52/
[Accessed March 2021].

Hildebrandt, M. & Gutwirth, S., 2008. *Profiling the European Citizen: Cross-Disciplinary Perspectives*. Netherlands: Springer.

Hill, R. K., 2016. What an Algorithm Is. *Philosophy&Technology,* Volume 29, p. 35–59.

Hood, C., 2006. Transparency in Historical Perspective. In: C. Hood, ed. *Transparency: The Key to Better Governance?.* New York: Oxford University Press, pp. 1-23.

Hughes, E., 1993. *Cypherpunk's Manifesto.* [Online]
Available at: https://www.activism.net/cypherpunk/manifesto.html
[Accessed March 2021].

KI Bundesverband, 2021. *Position Paper on EU-Regulation of Artificial Intelligence by the German AI Association,* s.l.: s.n.

Koene, A., 2019. *A governance framework for algorithmic accountability and transparency,* Brussels: European Parliament Scientific Foresight Unit (STOA).

Kuner, C., Svantesson, D., Cate, F. & Millard, O. L. C., 2017. Machine learning with personal data: is data protection law smart enough to meet the challenge?, ,. *International Data Privacy Law,* Vol. 7(No. 1).

Lessig, L., 1999. *Code and Other Laws of Cyberspace.* s.l.:Basic Books.

Lessig, L., 2000. Europe's 'me-too' patent law. *Financial Times,* pp. https://perma.cc/5H5H-KQBT.

Lessig, L., 2009. *Against Transparency.* [Online]
Available at: https://newrepublic.com/article/70097/against-transparency
[Accessed April 2021].

Lessig, L., 2020. *Introduction to Free Software, Free Society: The Selected Essays of Richard M. Stallman.* [Online]
Available at: https://www.gnu.org/philosophy/lessig-fsfs-intro.en.html

Maggiolino, M., 2019. *EU Trade Secrets Law and Algorithmic Transparency.* [Online]
Available at: Available at: https://ssrn.com/abstract=3363178 or http://dx.doi.org/10.2139/ssrn.3363178
[Accessed April 2021].

Mayer-Schönberger, V. & Cukier, K., 2013. *Big Data : A Revolution That Will Transform How We Live, Work and Think.* (John Murray 2013) ed. UK: Hachette.

McCarthy, J., 2021. *What is Artifficial Intelligence?.* [Online]
Available at: http://jmc.stanford.edu/artificial-intelligence/what-is-ai/index.html
[Accessed March 2021].

Mehrpouya, A. & Djelic, M.-L., 2014. *Transparency: From Enlightenment to Neoliberalism or When a Norm of Liberation Becomes a Tool of Governing,* s.l.: s.n.

Meyer, D., 2017. *How One European Smart City Is Giving Power Back To Its Citizens.* [Online]
Available at: https://www.alphr.com/technology/1006261/how-one-european-smart-city-is-giving-power-back-to-its-citizens
[Accessed April 2021].

Mittelstadt, B. & Wachter, S., 2019. A Right to Reasonable Inferences: re-thinking data protection law in the age of big data and AI. *Columbia Business Law Review,* Issue 2.

Morozov, E., 2013. *To Save Everything, Click Here: the Folly of Technological Solutionism.* New York: Public Affairs.

Nissenbaum, H., 1996. Accountability in a computerized society.. *Science and Engineering Ethics*, pp. 25-42.

OECD, 2017. *Recommendation of the Council on Open Government.* [Online]
Available at: https://www.oecd.org/gov/Recommendation-Open-Government-Approved-Council-141217.pdf
[Accessed March 2021].

O'Neil, C., 2016. *Weapons of Math Destruction: How big data increases inequality and threatens democracy.* New York: Crown Books.

Online Ethymology Dictionary, 2021. *Online Ethymology Dictionary.* [Online]
Available at: https://www.etymonline.com/search?q=transparent

Pasquale, F., 2015. *BLACK BOX ALGORITHMS: The Secret Algorithms That Control Money and Information.* London: Harward University Press.

Power, D. J., 2007. *A Brief History of Decision Support Systems.* [Online]
Available at: http://dssresources.com/history/dsshistory.html

Shalev-Schwartz, S. & Ben-David, S., 2014. *Understanding Machine Learning:From Theory to Algorithms.* ISBN 978-1-107-05713-5 Hardback ed. New York: Cambridge University Press.

Siraj, M., 2010. Exclusion of Private Sector from Freedom of Information Laws: Implications from a Human Rights Perspective. *Journal of Alternative Perspectives in the Social Sciences Volume 2, No 1,* pp. 211-226 .

Turing, A., 1950. Computing Machinery and Intelligence. *Mind, Volume LIX, Issue 236,* p. 433–460.

Tutt, A., 2017. An FDA for algorithms. *Administrative Law Review,* Issue 83.

Viktor Mayer-Schönberger and Kenneth Cukier, J. M. 2., 2013. *Big Data : A Revolution That Will Transform How We Live, Work and Think.* UK: Hachette Books.

Wachter, S., Mittelstadt, B. & Floridi, L., 2017. Why a right to explanation of automated decision- making does not exist in the General Data Protection Regulation. *International Data Privacy Law*, January.

Wachter, S., Mittelstadt, B. & Russell, C., 2018. Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR. *Harvard Journal of Law & Technology.*

Wagner, W. J., 1971. The Development of the Theory of the Right to Privacy in France. *Washington University Law Review*, pp. 45-70.

Warren, S. D. & Brandeis, L., 1890. The right to privacy. *Harvard Law Review 4,* p. 193–220.

Yu, H. & Robinson, D. G., 2012. The New Ambiguity of "Open Government". *UCLA Law Review Discourse*, pp. 180-208.