UNIVERSITA' COMMERCIALE 'LUIGI BOCCONI'

PhD SCHOOL

PhD program in Legal Studies

Cycle: XXXVII

Disciplinary Field (code): IUS08

The AI-driven public actor: new challenges to fundamental rights and the role of the private sector

Advisor: Oreste POLLICINO

PhD Thesis by

Flavia BAVETTA

ID number: 3001522

Year 2025

ABSTRACT

In the last few years, with the increasing use of technology, and especially of Artificial Intelligence (AI), new aspects in the exercise of the public actor's power can be observed. Specifically, as AI technologies begin to play a dominant role in the contemporary exercise of power, it becomes increasingly important to examine the phenomenology of a new kind of power and its unique challenges to constitutional principles.

On this merit, the public actor is making extensive use of technological and automated tools to reach fast and more efficient decisions. Considering that it has the privilege to acquire, *ex lege*, both *ex officio* or on the initiative of citizens, a wide availability of personal data and information, the development of AI technologies for public decision-making processes is now inevitable.

Therefore, the rise of a new paradigm can be observed, the so called 'AI-driven public actor', which has elements of absolute novelty.

In this context, this work has the aim of answering two major questions. First, which is the impact of the new paradigm of the 'AI-driven public actor' on the exercise of the citizens' fundamental rights, considering the guarantees that the European Union (EU) legal order is currently trying to ensure. Second, the thesis investigates whether the EU legal order provides for specific provisions to balance the role of the private actor in public decisions. Accordingly, the answers to these questions have important implications not only in terms of understanding the future of the application of fundamental rights, but also of the maintenance of democratic institutions, as it may depends on the ways in which new technologies are changing the relationship between the public actor, the private one and the civil society.

ACKNOWLEDGEMENTS

This has been a long and challenging project, which would not be the same, without the comments, help, and support of specific persons I owe the utmost gratitude towards.

Firstly, I wish to express my appreciation to my Supervisor, Professor Oreste Pollicino for his invaluable comments and academic support, especially at the most crucial moments of completing my thesis. His deep insights and knowledge made this work possible.

Secondly, I wish to thank Professor Marco Bassini for his guidance during my PhD visiting period at Tilburg University, during which he has provided constructive and precise remarks which have significantly improved my work.

Last but not least, I wish to thank colleagues, family and friends who have supported me through this work.

To everyone mentioned, and to everyone who has been part of this journey, thank you.

TABLE OF CONTENTS

CHAPTER I: INTRODUCTION IN A NUTSHELL

Contents: 1. Questions of this research; 2. The relevance of the chosen theme: Use of AI technologies by the EU public actor; 3. The aims of this research; 4. Methodology; 5. Definition of AI used in this research.

1. Questions of the research

This research deals with two core questions.

First, which is the impact of the Artificial Intelligence-driven public actor's paradigm on the exercise of citizens' fundamental rights? This core question requires answering the following sub-question: which are the limits and the guarantees that the European Union (EU) legal order is currently ensuring within this new paradigm? Second, does the public actor's procurement of private Artificial Intelligence (AI) gives to the private actor a specific role in public decisions? If yes, has the EU legal order established any provisions to balance the role of the private actor in public decisions? Moreover, this work – albeit briefly – researches whether the US legal order has enacted any provision dealing with these issues.

These two main questions are investigated on the basis of the following hypothesis.

In the last few years, with the increasing use of technology, and especially of AI, new characteristics in the exercise of the public actor's[1] power can be observed. Specifically, as AI begins to be used as an instrument in the contemporary exercise of power, it becomes increasingly important to examine the phenomenology of a new kind of power and its unique challenges to constitutional principles[2].

[1] In EU legal doctrine, a public actor generally refers to an entity or institution that performs public functions or exercises powers according to public law. Its key characteristics, *inter alia*, are: (i) being governed by public law, as public actors are typically established under national or EU law, and their functions are defined by statutory or regulatory provisions; (ii) the exercise of public authority, as these actors possess legal authority to make decisions or take actions that affect citizens; (iii) the use of public resources; (iv) accountability under EU law, as they are bound by EU principles such as proportionality, non-discrimination, and adherence to the EU Charter of Fundamental Rights. See *ex multis* P. Craig – G. De Búrca, *EU Law: Text, Cases, and Materials*, Oxford University Press, 2020; D. Chalmers - G. Davies – G. Monti, *European Union Law: Text and Materials,* Cambridge University Press, 2024.

[2] A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33.

The development of information and communication technologies and the consequent universal spread of tools and solutions such as information technologies, algorithms, or AI is giving a new direction in the way the public actor takes decisions that have an influence on the exercise of citizens' fundamental rights. On this merit, considering that it has the privilege to acquire *ex lege* – both *ex officio* or on the initiative of citizens – a wide availability of personal data and information, the development of AI technologies for public decision-making processes is now inevitable.

Therefore, the rise of a new paradigm, the so called 'AI-driven public actor'[3], can be observed with elements of absolute novelty with respect to (i) the instruments used to reach public relevant decisions; (ii) the need for finding a proper balance between the public actor's interests in using AI technologies and citizens' fundamental rights; and (iii) the strong technical and organizational interconnection between the public and private actors in the provision of public services[4]. More specifically, as described in the next chapters, public activity has been transformed by and through the application of new technologies, and particularly by AI[5]. Secondly, technological development invests economic and social relations to such an extent that existing rules need to be reassessed in order to confirm that the exercise of fundamental rights, such as privacy, personal data protection, and transparency of the public actor's action, is still guaranteed. Thirdly, a new and close interconnection exists between public and private actors considering that the public actor has neither the resources nor the expertise to develop such technologies, being then obliged to procure them from private companies.

[3] For the analysis of the development of this new paradigm, please see chapter II of this work.

[4] Other authors in the relevant literature have already analyzed some of the elements of the new AI-driven public actor paradigm presented in this work. For example, L. Torchia, *Lo stato digitale*, Il Mulino, 2023; O. Pollicino – G. De Gregorio, *Constitutional law in the algorithmic society*, Cambridge University Press, 2021, p. 3-24; A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33; A. Sanchez-Graells, *Resh (AI) ping Good Administration: Addressing the mass effects of public sector digitalization*, Laws, 2024, p. 9.

[5] A. Simoncini, *Amministrazione digitale algoritmica. Il quadro costituzionale, Il diritto dell'amministrazione pubblica digitale*, 2020.

Thus, the combination of these three aspects represents an important novelty that is investigated in this work.

2. The relevance of the chosen theme: the use of AI by the EU public actor

So far, the use of AI in public actor's decision processes is not only widespread but also supported through large investments. Indeed, the centrality of these new tools in the public sector and in government decision-making stems from the awareness of their enormous potential. Particularly, it is well known that AI allows the management of large databases to improve the work of public sector professionals, especially from the integration of internal and external databases, even if they include information of a different nature, and both quantitative and qualitative, to generate new results.

For this reason, different initiatives have emerged in the EU to promote AI development strategies in the public sector, as this technology has proven to be a powerful transformer on the European economy and competitiveness. Precisely, the Digital Agenda for Europe[6] is a set of policies that pretend to redesign the European telecoms sector and to achieve a data-driven public actor based on technology such as AI, blockchain, or big data. Moreover, the EU in its Coordinated Plan on Artificial Intelligence 2021[7] recognizes that AI applications can contribute to better public services, *i.e.*, by improving citizen-government interaction, enabling smarter analytical capabilities or improving efficiency across public-sector domains and supporting democratic processes. As such, it is recognized that the use of AI systems can bring benefits across all key public-sector activities.

Moreover, looking at national level, it should be noted that almost all Member States have enacted internal AI strategies including actions to stimulate the use of AI in the public

[6] European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions, A Digital Agenda for Europe* (May 19th, 2022).

[7] European Commission, *Coordinated plan on artificial intelligence 2021 review*, *Annexes to the Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Fostering a European approach to Artificial Intelligence*, (April 21st, 2021).

sector[8]. For instance, analyzing the Italian panorama, in the Strategic Program on Artificial Intelligence 2022-2024[9], the Italian government poses some critical goals in the use of AI in the public sector: on the one hand, Italy has to improve its internal processes and policies thanks to a responsible use of data and AI technology, and on the other hand, the government is committed to overseeing AI and mitigating its potential risks, especially to safeguard human rights and ensure an ethical deployment of AI[10].

Based on these premises, in order to understand the relevance of the chosen theme, it is worth providing some practical examples where the public actor uses AI systems to carry out certain activities of public interest. For this reason, the table below presents some of the current uses of AI by the public actor in the EU[11].

Table 1[12]

System	Country	Description
DG-Agri/ESA - The use of AI for satellite monitoring of European crops and compliance with CAP	EU	This use case concerns a pilot experiment in the field of the Common Agricultural Policy (CAP) aimed at satellite monitoring of European crops and compliance with agricultural subsidy rules. Currently, Member States are obliged to inspect five percent of subsidized crops on the ground in order to check compliance and prevent fraud.

[8] European Commission, *Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Fostering a European approach to Artificial Intelligence* (April 21st, 2021).

[9] Agenzia per l'Italia Digitale, *Strategia Italiana per l'intelligenza Artificiale 2024-2026*, (July 22nd, 2024).

[10] *Ibid*.

[11] The current inventory is by no means a complete overview of the use of AI by the public actor. In other words, the data are not and do not aim to be representative of the situation in Europe. The table has the only aim to show some of the practical use of AI by the public actor, as well as the possible benefits it can produce. However, some conclusions can be drawn from the mapping exercise. The first is the observation of limited information available on existing AI use cases within the EU administrations. It is striking that not only is this information is hardly available on the internet, but it is not even available to any centralized EU service. The establishment by Art. 71 of the AI Act of a centralized database within the Commission with the existing use cases in both the public and private sector is a positive step to overcome the current lack of unique and official information. The mapping exercise also revealed that there is considerable interest and growing use of AI tools by the EU administration itself. Moreover, it also confirms the importance of outsourcing in this area, and the limited capacities of the EU administration to develop its own AI systems. Furthermore, the use cases examined also show the great potential that AI can have for improving certain administrative functions, increasing their quality and effectiveness and not only reducing their cost.

[12] The examples provided are taken from different sources: (i) European Commission, Joint Research Centre, L. Tangi – A. R. Müller – M. Combetto et al., *Artificial Intelligence for interoperability in the European public sector – An exploratory study*, Publications Office of the European Union, (October 4th, 2023); (ii) L. Tangi – C. van Noordt – M. Combetto – D. Gattwinkel – F. Pignatelli, *AI Watch. European Landscape on the Use of Artificial Intelligence by the Public Sector* 2022; (iii) O. Mir Puigpelat, *Algorithms, Automation and Administrative Procedure at EU Level*, University of Luxembourg Law Research Paper, 2023; (iv) European Commission, *Public Sector Tech Watch latest dataset of selected cases*, 2024; (v) P. Bizzini, *The algorithm that blew up Italy's school system,* Algorithm Watch, (April 17th, 2023).

agricultural subsidy rules		The new system uses machine learning algorithms to improve the recognition accuracy of satellite images. It aims, among other things, to monitor all European fields, including those that are more difficult to access, and to reduce and optimize the number of field inspections, to the benefit of national administrations and farmers themselves, for whom the system can also make it easier to obtain subsidies. The system does not take automated decisions but merely issues alerts in cases of possible non-compliance. Such alerts are verified by humans through the review or zoom of images or, where appropriate, an on-site inspection, before a legal decision is taken to deny the requested subsidy or to reimburse the previously granted subsidy. Satellite monitoring can therefore form part of the complex procedures for the granting, control, and revocation of CAP subsidies.
EU-LISA - The use of AI for biometric recognition of persons at the EU's borders[13]	EU	This use-case refers to eu-LISA, the European agency responsible for the management of basic information systems for Member States' border and law enforcement authorities, such as the Schengen Information System (SIS II), the Visa Information System (VIS) and the asylum information system (Eurodac). Within this project, new information systems have been developed, such as the Entry/Exit System (EES), the European Travel Information and Authorisation System (ETIAS) and the European Criminal Records Information System - Third-Country Nationals (ECRIS-TCN), for their forthcoming entry into operation. AI is used in the first three systems and in the forthcoming EES and ECRIS-TCN for biometric identification and verification of people at EU borders and within Member States. All of them employ biometric matching systems, which use advanced machine learning algorithms to match facial images and fingerprints taken at the borders with those stored in these information systems. Each system has its own biometric matching service, but the companies developing the EES biometric system are also working on implementing a tool to enable simultaneous search and comparison of biometric data in all these information systems at the same time. These biometric matching systems are not developed by eu-LISA, but by private contractors on the basis of the technical specifications set by eu-LISA, which also tests their proper functioning.

[13] This use case turns out to be very interesting as a clear manifestation of the AI-driven public actor paradigm mentioned above. Indeed, eu-LISA is a system used by a public institution that has a strong impact on the fundamental rights of citizens (*i.e.*, freedom of movement, protection of personal rights, right to privacy) and that is provided by a private company that oversees the AI system.

		In any case, the systems that eu-LISA makes available to Member States are among the most advanced in the world and have a very high performance, superior to that of the most experienced border official. Their accuracy has increased tenfold since such systems began to be used by eu-LISA in 2014 and has been facilitated by the controlled environments in which they operate.
SAFERS - Structured Approaches for Forest fire Emergencies in Resilient Societies	Italy	SAFERS developed an innovative platform to improve the management of forest fires, as their impact has greatly increased due to climate change. With the aim of reducing the impact of future forest fires, SAFERS developed a new big data platform based on AI, coupled with other information systems, making use of different data sources such as the Copernicus Sentinel satellites. In addition to satellite data, the system analyses, in real time, data from *ad hoc* monitoring cameras, social media as well as mobile phones of professional users, volunteers and citizens who, through chatbots, may report the situation they are observing. SAFERS automatically detects fire events by combining different types of input data (*e.g.*, ground-based cameras, social networks or satellite data) that have different formats, validating, classifying and correcting information, and in this way, creating more accurate data. In addition, the system seeks to calculate measurement indices, such as the firewater index, a vulnerability indicator of the area to fires. Finally, it suggests actions to citizens on what to do and what not to do.
GPS algorithm for schools	Italy	The GPS algorithm for schools is a digitized and automatic procedure that has the aim to simplify and improve teacher recruitment in Italy. The algorithm rests on two rankings. The first is the *Graduatorie Provinciali per le Supplenze* (GPS) specifically designed for substitute teachers. The second is the *Graduatorie Provinciali a Esaurimento* (GAE) which is for teachers holding teaching qualifications. The algorithm evaluates teachers' CVs and cross-references their preferences for location and class with schools' vacancies. If there is a match, a provisional assignment is triggered, but the algorithm continues to assess other candidates. If it finds another matching candidate with a higher score, that second candidate moves into the lead. The process continues until the algorithm has assessed all potential matches and selected the best possible candidate for the role.
Procurement data quality and CPV improvement for TED data	Belgium	This AI-based solution aims at improving Tenders Electronic Daily (TED) data quality and to have better expenditure data and statistics. The Common Procurement Vocabulary (CPV) codes establish a single

		classification system for public procurement aimed at standardizing the references used by contracting authorities and entities to describe procurement contracts. These codes are introduced by a human who is trained through his or her own experience, knowledge, and environment, introducing involuntarily bias into the data. This project is in a proof-of-concept phase and has the objectives of suggesting to the user CPV codes from notice/procurement documents, elaborate statistics, and foreseeing which companies would be awarded with a contract in the future if a tender has certain characteristics. The system identifies the budgetary value in the procurement documentation and validates it. Moreover, it classifies the codes, and the result can be, for example, suggesting a ranking of CPV to the user with the objective of improving procurement data quality.
Automated transcripts to speed up judicial proceedings	Spain	Legal proceedings have a peculiarity when it comes to the drafting of judgments. Sometimes transcriptions of the proceedings are needed, especially when the judgment is appealed to higher instances. Given the recording of the proceedings, the video clip is examined to determine, for example, where exactly something was said or who was named. In the Basque Country, administration transcripts are made by manually reviewing the videos of all the sessions. Thus, it is not possible to easily search for words, phrases and generic entities across the video clip and there is not any correlation between the speech and the person who uttered it. Converting voice data into searchable text using automated transcription services may save significant time and create actionable value. In addition, if a court case is appealed to a higher instance, time is also saved by making it easier to find the exact points in the video in which specific information has been given. The solution provides speech-to-text functionality, digitizing the audio signal and keeping track of the audio and its transcription, as the legal value is still in the audio. The text is extracted in subtitle format and a web application is used to link the text with the concrete minute of the video/audio. Texts are saved and made available in a structured form.
OTT – decision support tool for consultants	Estonia	The decision-support system OTT is a tool that helps the Unemployment Insurance Fund's consultants quickly and effectively collect and analyze all the data about an unemployed person, taking into consideration the economy, and predicts the chances of them getting a new job. OTT applies AI trained and tested based on the last five years' unemployment data. Using the trained model and 60 different attributes and indicators, each unemployed

		person is evaluated, and their chances of finding a new job are calculated. The attributes are both about the person, *e.g.*, their education, previous job experience, right to benefits, health restrictions, and about the labor market, *e.g.*, the number and type of available positions in different regions and the number of newly unemployed people.

3. The aims of the research

AI systems are contributing to the introduction of new paths for innovation, thus producing positive effects for society as a whole. Technology is also an opportunity for shaping a new way of exercising the public function. Moreover, AI can provide better systems of enforcement of rules or improve the performance of public services.

Nonetheless, the domain of inscrutable AI characterizing contemporary society challenges the protection of fundamental rights and democratic values, encouraging lawmakers to find a regulatory framework balancing risk and innovation, considering the role and responsibilities of the public actor in the AI-driven age[14]. The challenges raised by AI technologies are not limited to freedom of expression, privacy, and data protection. Indeed, constitutional democracies are under pressure to ensure legal certainty and predictability of automated decision-making processes which collectively affect democratic values. Individuals are increasingly surrounded by technological systems that do not always ensure the possibility of understanding and controlling their underlying functioning[15]. Moreover, the role of the private actor, as well as the strict interconnection

[14] The term 'AI-driven age' refers to an era characterized by the pervasive integration of artificial intelligence (AI) into various aspects of society, industry, and daily life. This epoch signifies a paradigm shift in how decisions are made, processes are automated, and innovation is pursued, fundamentally altering human interaction with technology, labor markets, governance, and even cultural practices. See, *ex multis*, J. M. Balkin, *The Three Laws of Robotics in the Age of Big Data*, Ohio State Law Journal, 2017; H. Kissinger – E. Schmidt – D. P. Huttenlocher – S. Schouten, *The age of AI: and our human future*, Little, Brown and Company, 2021; O. Pollicino – G. De Gregorio, *Constitutional law in the algorithmic society*, Cambridge University Press, 2021.

[15] O. Pollicino – G. De Gregorio, *Constitutional law in the algorithmic society*, Cambridge University Press, 2021, p. 3-24.

with the public one for the provision of technologies and data, is changing the way public decisions are reached and public interest pursued.

Therefore, as declared in section 1 of this work, the issues analyzed investigate precisely the change in the way the public actor exercises its powers as a result of the adoption of AI systems. On this merit, chapter II of this work aims to understand how the transition to the new paradigm of the 'AI-driven public actor' has taken place, and to highlight which aspects require further investigation. Moreover, chapter III examines the critical constitutional aspects of the algorithmisation of the public actor, as well as whether a deterioration in the relationship between the state and the citizen can be witnessed, as individuals are increasingly subjected to fully automated decisions that have a strong impact on the exercise of their fundamental rights. For this reason, chapter III focuses on the European legal system in order to understand the direction the EU is taking in addressing this major new challenge. Additionally, in order to have a comprehensive view of the AI-driven public actor paradigm, it is essential to understand that the changes that have brought to its development are not just strictly linked to the new form of exercise of the public functions, but they are connected also to the role that the private actors are having in this transition. In this direction, chapter IV investigates the role of regulation in the field of AI and data sharing mechanism in order to discover if cooperative efforts between the public and private sector could lead to a balanced approach between risk and innovation. Lastly, being the use of AI by the public actor a spread phenomenon all over the world, this work in chapter V analyzes the AI Executive Order, even if it has been revoked by the current US administration, in order to understand the policy and legislative choices US could have adopted in this field.

On this matter, while recognizing the deep work already being carried out in the field of legal research in AI – as evident in the literature identified in this work –, the consolidated

analysis provided here hopes to offer further insights, given AI's increasingly widespread deployment and use, as well as its impacts on individuals and their fundamental rights[16], specifically in the public field.

4. Methodology

Regarding the methodology three assumptions are considered: (i) the approach of legal positivism is used as the theory that shapes the analysis from a methodological point of view; (ii) the definition of artificial intelligence corresponds with that provided in the Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 and, therefore, this work focuses its discussion solely on Machine Learning and, in particular, Deep Learning systems; (iii) the analysis is carried out from a constitutional point of view rather than from an administrative law point of view.

4.1 The approach used

Legal positivism is a legal theory that focuses on the separation of law as it is (*de facto*) from law as it ought to be (*de jure*). It views law as a system of rules created by recognized authorities, with its validity dependent on its source rather than its moral content. Positivism provides a structured framework for examining the legal frameworks governing technology without conflating them with ethical or policy considerations. More specifically, legal positivism can guide both objectives and methodology. Indeed, a positivist approach enables descriptive and analytical research goals. The descriptive aspect involves

[16] R. Rodrigues, *Legal and human rights issues of AI: Gaps, challenges and vulnerabilities*, Journal of Responsible Technology, 2020.

cataloguing and explaining the existing legal frameworks governing technology. The analytical aspect examines the coherence, consistency, and applicability of these laws without critiquing their moral or ethical implications. This methodology allows for a systematic study of laws that are directly relevant to this research topic.

As such, this methodological approach is used in this work, as the questions posed deal with fast-evolving regulatory landscapes and novel societal challenges. Indeed, this approach is particularly valuable in understanding how legal systems respond to new technological challenges, such as regulating AI in digital environments, as it is possible to carry out the analysis aseptically, critically examining the normative, jurisprudential and doctrinal sources on the topic.

However, this strength also introduces a limitation as positivism may not fully account for the fluid nature of technological advancements and the soft law instruments, such as industry guidelines, that influence the field. To address this limitation, the thesis acknowledges positivism's foundational role and supplement it with interdisciplinary insights or alternative perspectives in the next chapters.

Overall, legal positivism offers a rigorous and systematic methodology for analyzing the law in the field of new technologies. Its focus on the legal system as a coherent, autonomous structure, providing a clear framework for understanding the interplay between different legal norms and their role in regulating technology. While recognizing its limitations in addressing the dynamic and interdisciplinary nature of the field, the methodology remains a valuable foundation for the objectives and the structure of this research.

4.2 Definition of AI used in this research

Before looking at major legal challenges related to the use of AI by the public actor, it is worth setting the definition of AI that it is used within this work. As such, considering that this thesis mainly analyzes sources of European law it is inevitable to consider the definition provided in the Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonized rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (AI Act).

Specifically, according to art. 3(1), 'artificial intelligence system' (AI system) means:

> 'a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environment'.

Moreover, Recital 12 of the AI Act clarifies:

> 'The notion of 'AI system' in this Regulation should be clearly defined and should be closely aligned with the work of international organisations working on AI to ensure legal certainty, facilitate international convergence and wide acceptance, while providing the flexibility to accommodate the rapid technological developments in this field. Moreover, the definition should be based on key characteristics of AI systems that distinguish it from simpler traditional software systems or programming approaches and should not cover systems that are based on the rules defined solely by natural persons to automatically execute operations. A key characteristic of AI systems is their capability to infer. This capability to infer refers to the process of obtaining the outputs, such as

predictions, content, recommendations, or decisions, which can influence physical and virtual environments, and to a capability of AI systems to derive models or algorithms, or both, from inputs or data. The techniques that enable inference while building an AI system include machine learning approaches that learn from data how to achieve certain objectives, and logic and knowledge-based approaches that infer from encoded knowledge or symbolic representation of the task to be solved. The capacity of an AI system to infer transcends basic data processing by enabling learning, reasoning or modelling. The term 'machine-based' refers to the fact that AI systems run on machines. The reference to explicit or implicit objectives underscores that AI systems can operate according to explicit defined objectives or to implicit objectives. The objectives of the AI system may be different from the intended purpose of the AI system in a specific context. For the purposes of this Regulation, environments should be understood to be the contexts in which the AI systems operate, whereas outputs generated by the AI system reflect different functions performed by AI systems and include predictions, content, recommendations or decisions. AI systems are designed to operate with varying levels of autonomy, meaning that they have some degree of independence of actions from human involvement and of capabilities to operate without human intervention. The adaptiveness that an AI system could exhibit after deployment, refers to self-learning capabilities, allowing the system to change while in use. AI systems can be used on a stand-alone basis or as a component of a product, irrespective of whether the system is physically integrated into the product (embedded) or serves the functionality of the product without being integrated therein (non-embedded)'.

Hence, it is assumed that the notion of AI is related to computer systems capable of thinking, learning, collecting data and information from multiple sources, and acting

according to several objectives correlated to algorithms. The creation of these algorithms is a statistical, mathematical, and human process, which includes a large amount of data collection and analysis in different phases[17]. Particularly, the term of AI refers to systems that display intelligent behavior by analyzing the environment and taking actions to achieve specific goals with some degree of autonomy. The function of analyzing complex sets of data and taking decisions is strictly related to deep learning and machine learning. These utilities are the subfields or applications of AI, which deal with designing algorithms capable of educating machines by helping them recognize patterns and extract knowledge from previous cases[18]. An example of algorithms that apply deep learning are those used to calculate the risk of recidivism in criminal justice or in the identification of a specific number of aids considering the individuals they should be provided for. Deep learning is inspired by the functioning of neural networks in the human brain, and the data goes through different layers in which learning rules are applied[19]. Particularly, this system requires a sequence of instructions that specify the actions to be executed by the computer system, being able to provide solutions for complex problems[20]. This is precisely one of the keys to algorithms: their advantage lies in their ability to anticipate behaviors, guess trends, or witness plausibility.

At a first glance, algorithms seem like neutral technologies processing information which can lead to a new understanding of reality and predict future dynamics. Technically, algorithms, including AI technologies, are just methods to express results based on inputs

[17] C. Coglianese – D. Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era,* Georgetown Law Journal, 17, 2017, p. 147–223.
[18] A. Danish – S. Frimpong, *Artificial Intelligence, Machine Learning and Process Automation: Existing Knowledge Frontier and Way Forward for Mining Sector*, Artificial Intelligence Review, 53, 2020, p. 6025-6042.
[19] M. Bertolini – D. Mezzogori – M. Neroni – F. Zammori, *Machine Learning for Industrial Applications: A Comprehensive Literature Review*, Expert Systems with Applications, 175, 2021, p. 114-30.
[20] C. O'Neil, *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*, Crown Publishers, 2016.

made up of personal data and information[21]. According to Lustig and Nardi, algorithms 'are simultaneously a set of abstract instructions (logic) and possibilities for action (control)'[22]. It is then a sensible inference that algorithms are largely neutral in terms of the ways in which they process data. As such, they are seen to operate autonomously without the need for human intervention and have social and technical significance without being well understood by the general population of end users[23]. They are 'strictly rational concerns, marrying the certainties of mathematics with the objectivity of technology'[24,25]. However, this veil of neutrality falls before human fallacy. Processes operated by algorithms are value-laden since technologies are the result of human activities and determinations[26]. Indeed, the contribution of humans in the development of data processing standards causes the shift of personal interests and values from the human to the algorithmic realm. If, from a technical perspective, algorithms are instruments that extract value from data, then moving to the social perspective, such technologies constitute automated decision-making processes able to affect society and, thus, also impacting on constitutional values, precisely fundamental rights, and democratic values[27].

[21] T. Gillespie, *The Relevance of Algorithms*, in (eds.) T. Gillespie - P. J. Boczkowski - K. A. Foot, *Media Technologies: Essays on Communication, Materiality, and Society*, MIT Press, 2014, p. 167.
[22] C. Lustig - B. Nardi, *Algorithmic authority: The case of Bitcoin*, 48th Hawaii International Conference on the System Sciences, 2015, p. 743-752.
[23] M. Willson, *Algorithms (and the) everyday*, Information, Communication & Society, 2017, p. 137-150; N. Seaver, *Algorithms as culture: Some tactics for the ethnography of algorithmic systems*, Big Data & Society, 2017, p. 1-12; E. Rader - R. Gray, *Understanding user beliefs about algorithmic curation in the Facebook news feed*, Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2015, p. 173-182.
[24] N. Seaver, *Knowing algorithms*, Media in Transition, 2013, p. 1-12.
[25] H. Rosenbaum, *Algorithmic neutrality, algorithmic assemblages, and the lifeworld*, 2020.
[26] P. A. E. Brey – J. Soraker, *Philosophy of Computing and Information Technology*, Elsevier, 2009; N. Wiener, *The Human Use of Human Beings: Cybernetics and Society*, Da Capo Press, 1988.
[27] O. Pollicino – G. De Gregorio, *Constitutional law in the algorithmic society*, Cambridge University Press, 2021, p. 3-24.

4.3 The use of a constitutional perspective

Being the focus of this work on the use of AI systems by public actor, the analysis carried out is mainly from a constitutional law perspective rather than an administrative law one[28]. Indeed, as concerns with respect to the exercise of fundamental rights and individual freedoms are posed, the analysis has been oriented towards a constitutional point of view. Particularly, the fast-growing use of algorithms in the fields of justice, policing and public welfare could end in biased and erroneous decisions, boosting inequality, discrimination, unfair consequences, and undermining constitutional rights, such as privacy, freedom of expression, and equality. And these uses raise considerable concern not only for the specific policy area in which they are operated but also for society as a whole[29]. The domain of inscrutable algorithms characterizing contemporary society challenges the protection of fundamental rights and democratic values while encouraging lawmakers to find a regulatory framework balancing risk and innovation, considering the role and responsibilities of the public actor in the AI-driven age.

The challenges raised by AI technologies are not limited to freedom of expression, privacy, and data protection. Constitutional democracies are under pressure to ensure legal certainty and predictability of automated decision-making processes which can collectively affect democratic values. Individuals are increasingly surrounded by ubiquitous systems that do not always ensure the possibility of understanding and controlling their underlying technologies[30].

---

[28] On this matter, please see S. Ranchordas, *The Invisible Citizen in the Digital State: Administrative Law Meets Digital Constitutionalism*, in (eds.) J. De Poorter – C. Oirsouw – G. van der Schyff, *European Yearbook of Constitutional Law*, Tilburg Law School Research Paper, (December 24th, 2023), where the author states that 'Constitutional law and administrative law are inseparable partners in the protection of fundamental rights. Yet, in the digital state, where rights and access to public services are increasingly mediated by digital government platforms and automated systems, the role of administrative law has become invisible'.

[29] A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33.

[30] O. Pollicino – G. De Gregorio, *Constitutional law in the algorithmic society*, Cambridge University Press, 2021, p. 3-24.

Therefore, although the central role of administrative law in the use of AI systems by the public actor was considered during the drafting of this work, it was decided – and, therefore, preferred – to focus the analysis primarily on aspects of constitutional law, both in light of the importance of examining the guarantee of fundamental rights and due to the desire to focus on European law rather than on specific national laws.

CHAPTER II: THE 'AI-DRIVEN PUBLIC ACTOR': WHICH ARE THE CONSTITUTIONAL CHALLENGES?

Contents: 1. A new paradigm: the 'AI-driven public actor'; 2. Transformative role of AI in the public sector; 3. AI in public decisions challenges fundamental rights; 4. The interaction between public and private powers in the 'AI-driven public actor'; 4.1. The role of the private actor as enabler of the 'AI-driven public actor'; 4.2. Information sharing between private and public actors; 5. A path for further investigations.

1. A new paradigm: the 'AI-driven public actor'

In the last few years, with the increasing use of technology, and especially of AI, the rise of a new paradigm in the exercise of the public actor's power can be observed.

The characteristics of this new paradigm, the so called 'AI-driven public actor', have elements of absolute novelty with respect to (i) the instruments used to reach public relevant decisions; (ii) the need for finding a proper balance between the public actor's interests in using AI technologies and citizens' fundamental rights; and (iii) the strong technical and organizational interconnection between the public and private actors in the provision of public services[31]. The combination of these three aspects, thus, represents an important novelty that needs to be further investigated.

In this new context, in order to understand which major events have brought to the rise of this new paradigm, as well as to examine the interconnections between the aforementioned three new features in the exercise of public power, it is necessary to start from the very origin of constitutional theory dealing with the problem of power control.

[31] Other authors in the relevant literature have already analyzed some of the elements of the new 'AI-driven public actor' paradigm presented in this research. For example, *ex multis*: L. Torchia, *Lo stato digitale*, Il Mulino, 2023; O. Pollicino – G. De Gregorio, *Constitutional law in the algorithmic society*, Cambridge University Press, 2021, p. 3-24; A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33; A. Sanchez-Graells, *Resh (AI) ping Good Administration: Addressing the mass effects of public sector digitalization*, Laws, 2024, p. 9.

On this merit, scholars commonly consider constitutional law that part of the legal system whose function is to legally[32] delimit power[33]. In the modern sense, this discipline establishes rules or builds institutions capable of shielding personal freedoms from external constraints[34]. According to this idea, constitutionalism historically always adapted itself to power's features. Indeed, the protection of freedoms in constitutions has been shaped following the evolving character of the threats to those same freedoms.

In sum, at the beginning of the modern era, the power to be feared was the king's private force. The idea of sovereignty, which appeared at the end of the Middle Ages, had its roots in the physical and military strength of the very person of the sovereign. Sovereignty evoked an external power grounded on the monopoly (actual or potential) of the physical force used against individuals or communities (*e.g.*, military force or the force of law). Consequently, liberties were those dimensions of human life not subjected to that power (*e.g.*, *habeas corpus*)[35].

---

[32] Since this constitutional theory is part of the legal system this feature distinctively differentiates constitutional law from political philosophy or political sociology.

[33] G. Pino, *Il costituzionalismo dei diritti struttura e limiti del costituzionalismo contemporaneo*, il Mulino, 2017. On this merit, please consider that many scholars advocate this theory. For example, *ex multis*, M. Weber, *Economy and Society*, Bedminster Press, 1922, discusses the concept of legal-rational authority, emphasizing that modern states derive legitimacy from a system of laws, including constitutional frameworks that constrain and channel power. Moreover, N. Bobbio, *Futuro della democrazia*, 1984 argues for the importance of legal structures in defining the limits of power in democratic systems, emphasizing constitutionalism as a tool for preventing arbitrary rule. C. Schmitt, *Constitutional Theory*, Duke University Press, 1928, analyzes the role of the constitution in defining the boundaries of state power, particularly through the concept of 'constitutional order'. Additionally, H. Kelsen, *General Theory of Law and State*, Routledge, 1945, describes the constitution as the 'basic norm' (*Grundnorm*) that governs the hierarchy of legal rules and establishes the limits of governmental authority.

[34] On the concept of the 'rule of law', please consider, *ex multis*: A. Venn Dicey, *Introduction to the Study of the Law of the Constitution*, Roger E. Michener, 1885, where he articulates its key principles: (i) supremacy of regular law over arbitrary power; (ii) equality before the law; (iii) the role of the constitution in securing rights. Moreover, L. L. Fuller, *The Morality of Law, Indiana Law Journal*, 1964, where he identifies eight principles of legality (*e.g.*, generality, clarity, consistency) that underpin the rule of law. J. Raz, *The Authority of Law*, Oxford University Press, 1979, offers a procedural conception of the rule of law, emphasizing that it is not about moral content but about clear, stable, and prospective rules that guide behavior and limit discretion. H. L. A. Hart, *The Concept of Law*, Oxford University Press, 1961, discusses the rule of law as part of the broader legal system, focusing on the balance between primary and secondary rules that ensure legal predictability and governance.

[35] On this merit, *ex multis*: J. Locke, *Two Treatises of Government*, 1689, argues that individuals have natural rights to life, liberty, and property, which precede and limit governmental authority. Liberty, for Locke, is freedom from arbitrary power. Rousseau, *The Social Contract*, Penguin, 1762, distinguishes between legitimate political authority and the inviolable freedoms of individuals, which the state must respect. W. Blackstone, *Commentaries on the Laws of England*, 1765-1769, identifies individual liberties such as

As the offspring of the French and American Revolutions, the rule of law doctrine was the main legal tool 'invented' by constitutional theory to delimit the king's power and protect personal freedom and rights. As such, to be legitimate, any power has to be subject to the rule of law.

The other decisive turning point in the history of constitutionalism was World War II and the end of twentieth-century European totalitarian regimes. It may sound like a paradox, but those regimes showed that the legislative state, built on the supremacy of law and therefore exercising a legitimate power, can become another terrible threat to human freedom and dignity[36]. If the law itself has no limits, whenever it gives a right, it can withdraw it. This practice is the inhuman history of some European twentieth-century states that cancelled human dignity through the law.

With the end of World War II, a demolition process of those regimes began, and learning from the American constitutional experience, Europe transformed flexible constitutions – until then, mere ordinary laws – into rigid constitutions, which are effectively the supreme law of the land. In this new scenario, the power that instils fear is no longer the king's private prerogative. The new force is the public power of state laws, and the constitutional

---

personal security, personal liberty, and private property as fundamental rights protected from government encroachment by laws like habeas corpus. A. Venn Dicey, *Introduction to the Study of the Law of the Constitution*, Roger E. Michener, 1885, highlights habeas corpus as a cornerstone of English constitutionalism, ensuring that individual liberty is protected from arbitrary detention. L. L. Fuller, *The Morality of Law, Indiana Law Journal*, 1964, argues that legal systems must safeguard individual freedoms by adhering to principles that ensure the predictability and fairness of laws, implicitly supporting the notion of liberties as zones free from arbitrary power. J. Raz, *The Authority of Law*, Oxford University Press, 1979, emphasizes that individual autonomy is a central goal of the rule of law, requiring the state to respect areas of life where individuals exercise independent judgment.

[36] On this merit, *ex multis*: C. Schmitt, *Constitutional Theory*, Duke University Press, 1928, critiques liberal constitutionalism, arguing that legal norms could be manipulated under a legislative state to legitimize extraordinary powers, a view often tied to his controversial support for authoritarianism. H. Kelsen, *General Theory of Law and State*, Routledge, 1945, acknowledges the dangers of legislative dominance when unrestrained by constitutional safeguards or democratic accountability. N. Bobbio, *Liberalism and Democracy*, 1990, discusses how formal legal systems could devolve into instruments of tyranny if they lacked substantive protections for human rights and democratic principles. G. Sartori, *Comparative Constitutional Engineering*, 1994, examines how legal and legislative frameworks in authoritarian regimes can undermine constitutionalism by prioritizing state power over individual liberties. L. L. Fuller, *The Morality of Law, Indiana Law Journal*, 1964 argues that law must meet substantive moral criteria to prevent legislative systems from becoming instruments of oppression.

tool intended to effectively regulate that power is vested in the new rigid constitution: a superior law, stronger than ordinary statutes and, thus, truly able to protect freedom, at least apparently, even against legal acts.

However, with the turn of the twenty-first century, the advent of new digital technologies has provided an unprecedented means of limiting and directing human freedom. In this context, the rise of a new way of exercising power both from public and private actors can be witnessed[37].

Indeed, this technological power is at the origin of 'platform capitalism', which is a vast economic transformation induced by the exponentially fast-growing markets of Internet-related goods and services – for example, smart devices (Apple, Samsung, Huawei, Xiaomi), web-search engines (Google), social media corporations (Facebook, Instagram, Twitter), cloud service providers (Amazon, Microsoft, Google), e-commerce companies (Amazon, Netflix), and social platforms (Zoom, Cisco Webex). These 'moligopolists' are not only creating communities and benefitting from network effects generated by users' transactions, but they also develop a *de facto* political authority and influence once reserved for legal and political institutions. More importantly, they are taking on configurations that are increasingly similar to the state and other public authorities. Their structure reflects a fundamental shift in the political and legal systems of Western democracies – what has been called a new type of 'functional sovereignty'[38].

Moreover, the development of information and communication technologies and the consequent universal spread of innovative tools and solutions is giving a new direction in the way the public actor ensures the exercise of citizens' fundamental rights. On this merit, the public actor is making extensive use of technological and automated tools to reach

[37] A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33.
[38] *Ibid*, p. 27-33.

fast and more efficient decisions. Considering that it has the privilege to acquire *ex lege*, both *ex officio* or on the initiative of citizens, a wide availability of personal data and information, the development of AI technologies for public decision-making processes is now inevitable. Particularly, the most advanced systems of AI based on the analysis of data are variously and widely used to support public decision-making in order to, for instance, facilitate automated control of social funds, infer insights from large datasets in the health sector, preventing and detecting school absence and early leaving[39]. Therefore, a transformation into an AI-driven public actor can be witnessed.

In this context, from a constitutional point of view, the most problematic phenomenon is not the development of the general digitization of the public actor (and of its public administration)[40], but rather its evolution to the 'AI-driven public actor', as the growing use of AI systems for the purposes of decision-making in the field of public services is sensitively transforming the way the public action is carried out[41], as well as the procedural guarantees that citizens can actually exercise. To understand the issues related to this change, it is necessary to start from the most recent evolution of digital technology, due to the advent of the so called Internet of Things, which is rapidly transforming into the Internet of Everything, also thanks to a vast amounts of data (the so-called big data) which derive from the widespread diffusion of commonly used devices (phones, credit cards,

[39] See also T. S. Gesk – M. Leyer, *Artificial Intelligence in public services: When and why citizens accept its usage*, Government Information Quarterly, 2022, p. 1-12; H. Gieskea – I. Van Meerkerk – A. Van Buuren, *The Impact of Innovation and Optimization on Public Sector Performance: Testing the Contribution of Connective, Ambidextrous, and Learning Capabilities*, Public Performance and Management Review, 2019, p. 432-460.

[40] In relation to the digitalization of the public administration, see *ex multis*: M. Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*, Edward Elgar Publishing, 2015; C. Coglianese, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, Georgetown Law Journal, 2017; L. Edwards, *Data Protection and Legal Regulation of AI in Europe*, Computer Law & Security Review, 2018.

[41] G. Schneider, *The Algorithmic Governance of Administrative Decision-Making: Towards an Integrated European Framework for Public Accountability*, Eurojus, 2019, p. 134-148.

home appliances, cars) that, precisely, continuously produce and transmit data about people and things[42].

Consequently, the exercise of public power in an AI-driven environment is taking on new tools and means based primarily on data, not territory, and specifically on harvesting data that individuals produce and share by using digital platforms or through surveillance by cameras and devices set up in phones, computers, cars, home appliances, and so forth. Controlling this vast amount of data, which potentially shows attitudes, behaviors, and actions of billions of people, empowers the public actor to determine whether and on what conditions individuals may access public or private services and goods.

Therefore, the increasing reliance on digital platforms for communication and for providing even the most basic goods and services is diminishing the space for accessing resources offline and without the intermediation of a public or private actor[43]. In this context, the way the traditional public actor exercises its powers[44] has been overtaken by the concept of an AI-driven public actor, which points out more precisely this novelty. As such, while the public actor continues to perform all the functions and tasks it previously assumed, it has three new features. Firstly, public activity has been transformed by and through the application of new technologies, and particularly AI, in order to re-articulate and reorganize public functions. Indeed, algorithms are not used only to draft an administrative act, to preserve it or to transmit it, but they are used to determine its content in order to decide and orient public decisions[45]. Secondly, technological development invests in economic

---

[42] A. Simoncini, *Amministrazione digitale algoritmica. Il quadro costituzionale,* Il diritto dell'amministrazione pubblica digitale*,* Torino, 2020.

[43] K. Pistor, *Statehood in the Digital Age*, Constellations, 2020, p. 3.

[44] According to S. Junginger, *Transforming Public Services by Design: Re-Orienting Policies, Organizations and Services around People*, Routledge, 2016, the purpose of public organizations is to mediate the relationships between government and citizens and make positive contributions to society, by providing their services to citizens as well as instruments for implementing public policies. The public sector has to abide by the social contract that grants legitimacy to its pursuit to maximize public value for all. See, J.J. Rousseau, *The social contract*, Penguin, 1762.

[45] A. Simoncini, *Amministrazione digitale algoritmica. Il quadro costituzionale,* Il diritto dell'amministrazione pubblica digitale, Torino, 2020.

and social relations to such an extent that existing rules are often unsuitable or obsolete to guarantee the exercise of fundamental rights (*i.e.*, privacy, personal data protection, transparency of the public administration's action). Thirdly, a new and close interconnection exists between public and private actors considering that the public actor – and public administrations in particular – have neither the resources nor the expertise to develop such technologies[46]. Thus, analyzing the features of AI-driven public actor means studying how AI technologies produce a change of paradigm in the public decision-making processes. The difficulty of this analysis derives from the fact that the main force of algorithms is their practical convenience, so their interference with freedoms could be not perceived as an external constraint or a disturbing power. However, while producing highly effective practical outcomes, algorithmic decisions within the new paradigm of the AI-driven public power could undermine procedural and substantive guarantees related to democracy and the rule of law[47].

Therefore, considering the three new features of this new paradigm, in this chapter the major constitutional concerns related to the use of AI by the public actor will be analyzed. Particularly, in section 2, the role of AI in the exercise of public power is examined. Moreover, in section 3 the challenges that AI poses to fundamental rights are briefly presented. Lastly, in section 4, the interactions between public and private powers in this new paradigm are taken into account.

[46] Please note that, globally there are several proposals and attempts to regulate the use of artificial intelligence. For example, in the EU the AI Act has been recently enacted. Moreover, in the United States, the Executive order on the safe, secure, and trustworthy development and use of artificial intelligence was issued on October 30th, 2023. Additionally, on September 9th, 2024, China's National Technical Committee 260 on Cybersecurity released the first version of its AI Safety Governance Framework (the Framework), which was formulated to implement the Global AI Governance Initiative.

[47] A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33.

2. Transformative role of AI in the public sector

As defined in section 1, the growing use of algorithms for the purposes of decision-making in the field of public services is sensitively transforming the way public action is carried out[48]. According to Martin Painter's analysis of public sector reform, the last fifteen years have seen a 'paradigmatic realignment of state, markets and civil society' which has been galvanized by open data initiatives, self-governance, social enterprise, decentralization, and network governance[49]. Data-driven techniques have contributed to this shift. Moreover, also the European Commission has recognized the potential of AI in the public sector, and it has included a specific section in its White Paper on AI[50] to promote its adoption. The plan mainly focuses on four key areas: (i) increasing investment; (ii) making more data available[51]; (iii) fostering talent; and (iv) ensuring trust. Also, in the 2021 Review of the Coordinated Plan on Artificial Intelligence[52], the public sector is pointed out as a 'trailblazer for using AI', and the Commission remarks that it is essential that public administrations, hospitals, utility and transport services, financial supervisors, and other areas of public interest rapidly begin to deploy products and services that rely on AI in their activities[53].

As such, the current AI-driven era is based on an information economy and the evolution in data management has fostered the adoption of e-government strategies. From the incorporation of information computer technologies (ICTs) to the algorithmisation of some

[48] G. Schneider, *The Algorithmic Governance of Administrative Decision-Making: Towards an Integrated European Framework for Public Accountability*, Eurojus, 2019, p. 134-148.
[49] C. Painter, *The UK Coalition government: Constructing public service reform narratives*, Public Policy & Administration, 28, 2013, p. 3–20.
[50] European Commission, *White Paper on Artificial Intelligence - A European approach to excellence and trust*, (February 19th, 2020).
[51] Please, see also section 4 of this chapter and chapter IV.
[52] European Commission, *Coordinated plan on artificial intelligence 2021 review*, *Annexes to the Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Fostering a European approach to Artificial Intelligence*, (April 21st, 2021).
[53] European Commission, Joint Research Centre, L. Tangi – A. R. Müller – M. Combetto et al., *Artificial Intelligence for interoperability in the European public sector – An exploratory study*, Publications Office of the European Union, (October 4th, 2023).

services, the public sector has sought to modernize and improve the way decisions are made and services delivered in order to keep pace with private sector developments. This trend has seen governments attempt to transition from a programmatic model of service delivery to a citizen-focused model. Alongside this, service delivery has had to advance from an approach focused solely on service quality to a model that emphasizes the delivery of better outcomes more efficiently. In this renovated and more technological context, the benefits generated from AI's sophisticated data analytics in government are vast. AI capabilities clearly outstrip human ones in certain areas, such as data processing ability, and improvement in service delivery outcomes have been demonstrated[54]. It can help organizations work more effectively, reduce costs, and improve the quality of services provided to citizens. By using AI to partially automate tasks that would otherwise require human intervention, public sector organizations can free up resources to focus on more pressing issues. AI can also help public sector organizations make more informed decisions by analyzing large amounts of data and identifying patterns and trends that would be difficult to detect manually. Moreover, AI enables public sector organizations to work more efficiently, make better decisions, and deliver better services. For example, AI-powered chatbots can provide faster and more accurate responses to citizens' queries, freeing up staff to focus on more complex tasks. Predictive analytics can help healthcare providers predict and prevent health issues, resulting in better outcomes for patients. AI can also improve public procurement by automating repetitive tasks, identifying potential suppliers, and analyzing contracts for compliance[55]. Specifically, the public sectors which AI can be applied to are vast and open to clear efficiency improvement of public actor's

[54] Centre for Public Impact, *Destination unknown: Exploring the impact of Artificial Intelligence on Government, Working Paper*, 2017.
[55] European Commission, Joint Research Centre, L. Tangi – A. R. Müller – M. Combetto et al., *Artificial Intelligence for interoperability in the European public sector – An exploratory study*, Publications Office of the European Union, (October 4th, 2023).

duties[56]. Particularly, ranging from welfare and criminal justice, to healthcare, national security and beyond, governments are increasingly relying on algorithms to automate their decision-making processes[57]. These patterns have been amplified with the increasing employment of algorithms as processing infrastructures of public actor's collected data. Therefore, through algorithms processing, data is not any more static evidence, working as a support to public decisions carried out by public officials, but in those sectors where algorithms could be employed, data is becoming the center and the source of the decision-making processes[58]. Indeed, the increasing quantitative and qualitative importance of algorithms for the purposes of decision making in various fields of the public sector is transforming algorithms into outright governance tools: algorithms are employed as a means for authorities to potentially manage individual behavior, allocate resources and make public actor's procedures faster.

Thus, it appears that, similarly to the private sector, also the public one is being overtaken by a new form of AI-driven governance[59]. Through the 'algorithmisation' of the public action[60], the structure of public affairs is shifting to regard also as technical problems that

[56] For few practical examples, please see table 1, in chapter I.

[57] M. Zalnieriute - L. B. Crawford - J. Boughey - L. B. Moses - S. Logan, *The Cambridge Handbook on the Law of Algorithms*, Cambridge University Press, 2019, p. 30.

[58] K. Yeung, *Algorithmic Regulation: a Critical Interrogation*, Regulation & Governance, 2017, p. 20.

[59] It is undeniable that any aspect related to individuals' everyday life has been strongly influenced by digitization through the establishment of a system in which algorithms are used to collect, store, and organize citizens' data in order to make all kinds of decisions. The society is, therefore, subject to a new form of governance, *i.e.,* algorithmic governance, that sees AI systems as decision-making tools. In this context, it should be noted that the term 'algorithmic regulation' was first coined by T. O'Reilly, *Open Data and Algorithmic Regulation*, in (eds.) B. Goldstein - L. Dyson, 2013, p. 289-300. Then the idea of algorithmic society has been explained, *ex multis*, by J. Danaher, in *The Threat of Algocracy: Reality, Resistance and Accommodation*, Philosophy & Technology, 2016, p. 245–68.

[60] C. Coglianese - D. Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, The Georgetown Law Journal, Georgetown Law Journal, 2017, p. 147–223. In this context, it is interesting to consider Coglianese's analysis as he makes a step further affirming that: 'The transition to this online interaction with government over the last quarter-century portends what will likely be a deeper and wider technological transformation of governmental processes over the next quarter-century. Moving beyond the digitization of front-end communication with government, the future will likely feature the more extensive automation of back-end decision-making, which today still often remains firmly in the discretion of human officials. But we are perhaps only a few decades away from an administrative state that will operate on the basis of automated systems built with machine learning algorithms, much like important aspects of the private sector increasingly will. This will lead to an administrative state characterized by what I have elsewhere called algorithmic adjudication and robotic rulemaking'.

need to be addressed through technical solutions. Moreover, with algorithms becoming the engines of public affairs' management, and, thus, exerting control over society, a new form of technocratic public governance emerges, aiming at providing decisions on the basis of data-driven models. Ultimately, the data-driven models are creating a new form of 'technological determinism'[61] given by algorithms' predictions triggering public action. Indeed, algorithms rely on 'actuarial predictions' given by the correlations between the features or characteristics drawn from the data.

As framed in these terms, the AI-driven public governance is drastically overturning the traditional ways of managing public affairs[62]. In the current technological environment, there is a risk that the human evaluation factor is completely replaced by machine-driven calculations[63]. In this context, the global trend of using data analytics and AI tools throughout all levels of governments warrants a more engaging and informed approach to tackle the possible lack of *i.e.*, legality, accountability, and transparency in the use of such technology, and ultimately the transformation of fundamental rights as principles with no practical relevance and scope of application. Moreover, the adoption of AI in the public sector poses several challenges, ranging from technological to organizational and legal perspectives, and it is hindered or driven by a number of factors, such as the existence of

[61] On the idea of 'technological determinism' see, *ex multis*: T. P. Hughes, *Technological Momentum*, in (eds.) M. R. Smith – L. Marx, *Does Technology Drive History? The Dilemma of Technological Determinism*, MIT Press, 1994, p. 112; J. E. Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism*, Oxford University Press, 2019; M. Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*, Edward Elgar Publishing, 2015.

[62] This new paradigm has been also identified in terms of 'Digital constitutionalism' which asks broader questions that require to rethink power asymmetries between government and citizens, the relationship between private technology companies and the pursuit of the public interest, the safeguard of fundamental rights such as freedom of speech, and the protection of public values in the digital state. E. Celeste, *Digital constitutionalism: a new systematic theorisation*, International Review of Law, Computers & Technology, 2019, p. 76-99 defines digital constitutionalism as an 'ideology that adapts the values of contemporary constitutionalism to the digital society'. Please see also S. Ranchordas, *The Invisible Citizen in the Digital State: Administrative Law Meets Digital Constitutionalism*, in (eds.) J. De Poorter – C. Oirsouw – G. van der Schyff, *European Yearbook of Constitutional Law*, Tilburg Law School Research Paper, (December 24th, 2023); C. Coglianese – D. Ben – M. Lavi, *AI in Adjudication and Administration*, 2021, p. 2118; Y. Meneceur, *Artificial Intelligence, Public Administration, and the Rule of Law*, in (eds.) M. Suksi, *The Rule of Law and Automated Decision-Making*, Springer, 2023.

[63] G. Schneider, *The Algorithmic Governance of Administrative Decision-Making: Towards an Integrated European Framework for Public Accountability*, Eurojus, 2019, p. 134-148.

a suitable infrastructure, the right competences and skills as well as a fruitful collaboration with the supplier, as governments across the world are giving increased attention to the potentiality of AI[64].

As such, it is worth analyzing how the use of AI technologies by the public actor can affect the guarantee of citizens' fundamental rights, as public decisions are deeply relevant to civil society. Moreover, the deep interconnection between private and public actors in the technological sectors triggers the autonomy and independence of the public actor's decisions processing.

Accordingly, investigating this set of problems is crucial to understanding whether it is acceptable from a constitutional point of view that AI systems take part to the public decisions processes.

3. AI in public decisions challenges fundamental rights

Within the new AI-driven public actor paradigm, certain concerns with respect to the exercise of fundamental rights and individual freedoms[65] are posed. Particularly, the fast-growing use of algorithms in the fields of justice, policing and public welfare could end in biased and erroneous decisions, boosting inequality, discrimination, unfair consequences, and undermining constitutional rights, such as privacy, freedom of expression, and equality. And these uses raise considerable concern not only for the specific policy area

[64] European Commission, Joint Research Centre, L. Tangi – A. R. Müller – M. Combetto et al., *Artificial Intelligence for interoperability in the European public sector – An exploratory study*, Publications Office of the European Union, (October 4th, 2023).

[65] On the risks to individual freedoms, the best known and most circumstantial accusation can be found in S. Zuboff, *The Age of Surveillance Capitalism. The Fight for a Human Future and the New Frontier of Power*, Public Affairs, 2019, according to 'The vast lawless regions of the digital became the landscape in which companies and governments ruthlessly battle for information dominance, reenacting earlier epochs of invasion, conquest, and empire building in the physical world' and 'Surveillance capitalism originates in an even more startling mental invention, declaring private human experience as free raw material for translation into production and sales. Once private human experience is claimed for the market, it is translated into behavioral data for computational production'.

in which they are operated but also for society as a whole[66]. More specifically, alarms are primarily related to the inherent functionalities of AI, as they can be in violation of constitutional principles of legality, legitimacy, transparency, impartiality, efficiency, accountability, and good administration.

For example, regarding the duty to guarantee transparency, it should be noted that the use of AI has to deal with the existence of the black box problem[67], opacity and absence of legal reasoning. Indeed, since algorithms used in public decisions end up in determining public policy directions, they must be comprehensible for ensuring the right of access. Furthermore, in terms of impartiality, it should be considered that algorithms could make biased decisions. An algorithm's design and functionality could reflect the values of its designer, as well as the biased patterns based on the set of personal data chosen to train it. Moreover, another concern is related to the possible discriminatory decision taken by the AI system on certain characteristics such as gender, race, sexuality, or religious belief, as one group might be disproportionately harmed or benefited. On this merit, there is ample literature[68] with regard to forms of so-called institutional discrimination, particularly regarding racial profiles. In addition, the use of AI can pose some concerns in relation to the guarantee of the right to good administration, as it includes, but is not limited to, the right of individuals to have access to their file and the obligation of any public authority to give sufficient reasons for its decisions. Particularly, a question arises on how to ensure

[66] A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33.

[67] F. Pasquale, *The black box society: the secret algorithms that control money and information*, Harvard University Press, 2015. For purposes of this chapter, black box AI refers to any natural language processing, machine learning, textual analysis, or similar software which uses data not accessible to the data subject, and/or which deploys algorithms which are either similarly inaccessible, or so complex that they cannot be reduced to a series of rules and rule applications comprehensible to the data subject.

[68] Please see, *ex multis*: K. Crawford – D. Roel – T. Dryer – G. Fried – B. Green – E. Kaziunas – A. Kak – V. Mathur – E. McElroy – A. N. Sánchez – D. Raji – J. L. Rankin – R. Richardson – J. Schultz – S. Myers West – M. Whittaker, *AI Now 2019 Report*, AI Now Institute, 2019; C. Intahchomphoo – O. E. Gundersen, *Artificial Intelligence and Race: a Systematic Review*, Legal Information Management, 2020, p. 74-84; L. Koen – M. Kgomotso, *Artificial Intelligence and Racial Discrimination*, in (eds.) J. Temperman – A. Quintavalla, Artificial Intelligence and Human Rights, 2023.

that the potentially huge number of individuals have access to their files and specifically, to personal data used to train AI systems as well as to take specific decisions. Another question is how to make sure that public authorities always give sufficient reasons when the operation of AI-driven technologies cannot be fully explained due to their inherent opacity and complexity. Lastly, the use of AI can pose some concerns when it is necessary to address accountability of the public actor arising from misuse of AI and software fallibility. In this scenario, the very strong interconnection with the private sector – as it is the major provider of advanced technologies – poses the problem to assess who would be liable in case of misfunction of the AI. In addition to the above, the absence of strict regulation in the cybersecurity field can pose major concerns considering the exposure of citizens' life and the impact that automated decision processes can have on them. Particularly, AI systems are increasingly subject to manipulation, exposing citizens' personal data and information to major threats. Indeed, it is necessary to reflect how every device connected to a network is a potential channel for the existing cyber threats (*i.e.*, ransomware attacks, Distributed Denial of Services) to enter the public actor's systems. This expands the scope of possible vulnerabilities, which relate to the possibility that these rights cannot be exercised at all, since the services provided by the public actor may be altered or may be deprived of any functionality. Therefore, the use of highly sophisticated technology must always be accompanied by a high level of cybersecurity of its systems in order to ensure public actor's smooth functioning. This has to become a priority, redefining the centrality of cybersecurity also in the public field.

However, fundamental rights do not exhaust the threats which these technologies raise for constitutional democracies. The spread of automated decision-making can also challenge democratic systems due to its impact on public discourse and the impossibility of understanding decisions that are made by automated systems affecting individual rights

and freedoms[69]. This is evident when focusing on how information flows online and on the characteristics of the public sphere, which is increasingly personalized rather than plural. Likewise, the field of data is even more compelling due to the ability of data controllers to affect users' rights to privacy and data protection by implementing technologies the transparency and accountability of which cannot be ensured. The possibility of obtaining financing and insurance or the likelihood of a potential crime are only some examples of the efficient answers which automated decision-making systems can provide and of how such technologies can affect individuals' autonomy[70].

In light of these problems, the research in chapter III analyzes whether any legal safeguard and procedural guarantee exists for citizens in order to protect the exercise of their fundamental rights.

[69] On this matter, please see the reflections of D. Tiberiu – Y. Lupu, *Digital authoritarianism and the future of human rights*, International Organization, 2021, p. 991-1017, where they argue that technology can also have a negative impact on human rights. Without denying the beneficial effects of digital technologies to opposition groups, many argue that technological advancements also empower authoritarian governments by facilitating preventive repression. Preventive repression is often the first and most important line of defense for authoritarians. Preventing potential opposition groups from organizing and being publicly heard has long been an essential feature of authoritarian governments, from the creation of Fouche's secret police to augment Napoleon's rule to Metternich's use of political police to buttress the Habsburgs' power. Preventive repression can have multiple effects on dissenting groups. It can raise the cost of mobilizing to challenge the state by disrupting the dissenting group, cutting off their communication, making gathering more difficult, and restricting access to resources. Technological innovation provides authoritarian governments a wider set of tools with which to conduct these activities, thus lowering the cost of preventive repression. Contemporary authoritarian governments have consistently and effectively used technological tools to abuse human rights and further their own anti-democratic ends. Recently, for example, such governments have used voice recognition to scan mobile networks, tracked citizens' movement using GPS, read emails and text messages in order to monitor dissident groups and selectively censor information, and used malware and spyware to secretly turn on webcams built into personal laptops and microphones in cell phones. The Internet, in particular, facilitates the use of many tools that benefit such governments, including sophisticated digital monitoring. New technology thus impacts both (a) the government's ability to preventively stifle opposition groups' attempts to mobilize public protest and (b) opposition groups' ability to mobilize dissent. On this merit, see also S. Feldstein, *The Rise of Digital Repression*, Oxford University Press, 2021.

[70] O. Pollicino – G. De Gregorio, *Constitutional law in the algorithmic society*, Cambridge University Press, 2021, p. 3-24.

4. The interaction between public and private actors in the 'AI-driven public actor'

In order to have a comprehensive view of the AI-driven public actor paradigm, it is essential to understand that the changes that have brought to its development are not just strictly linked to the new form of exercise of fundamental rights, but they are connected also to the role that the private actors are having in this transition.

Two aspects of this strong interconnection between public and private actors can be observed. First, more frequently private actors hold the know-how for the creation and use of sophisticated technologies such as AI. Therefore, the public actor has to rely on the provision of technologies by the private sector which, usually, is not open to sharing proprietaries information about the functioning of the AI systems. Second, in order to enhance transparency between the public and the private actor, as well as in order to create positive externalities, the EU is fostering a strategy on data sharing between them. On this merit, the European Commission in the European strategy for data[71] has stated that:

> 'The value of data lies in its use and re-use. Currently there is not enough data available for innovative re-use, including for the development of artificial intelligence. The issues can be grouped according to who is the data holder and who is the data user, but also depend on the nature of data involved (*i.e.*, personal data, non-personal data, or mixed data-sets combining the two). Several of the issues concern the availability of data for the public good'[72].

[71] European Commission, *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*, (February 19th, 2020).
[72] *Ibid*.

Therefore, in these last years several regulations[73] have been enacted in order to create mechanisms of data sharing.

In this context, it is interesting to examine how the private sector has taken part in the 'algorithmisation' of the public one due to the technological dependency that exists between these two actors. In addition, the regulatory phenomenon concerning data sharing obligations between the public and the private actors is examined.

4.1 The role of the private actor as enabler of the 'AI-driven public actor'

In order to understand the role of the private actors as technological enablers of the 'AI-driven public actor', it is necessary to describe the phenomenon of the development of big tech corporations, as it is recent enough to allow a brief summary of the main steps.

On this merit, the large companies that now control the technology market – and beyond – have been able to develop in a very favorable regulatory regime, characterized by few constraints and many incentives.

First, a poor regulatory environment was a decisive condition for the creation of innovative services, the transformation of startups into platforms of enormous global scale, as well as for the still ongoing growth of investment and market. Indeed, the belief that regulation would have hindered and, perhaps, made innovation impossible to spread, as well as the idea that a more interventionist approach would have entailed more disadvantages than advantages, has permitted the exclusion of new technological services from the firt regulatory attempt[74].

[73] Particularly, the Data Act, the Data governance Act and Open Data Directive have been enacted by the EU also to foster the communication of data and information between the private and public sector.

[74] Particularly, during the Clinton presidency, the US government decided to not apply the 'information services' discipline to the new technological platforms even if it was already in place for telecommunication operators.

Moreover, the transformation from the status of startups to large companies has been based also on the possibility and ability to use the Internet as a single, undifferentiated tool without access costs. This is exceptional for two different reasons. First, these companies operate having a previously unprecedented global dimension, thanks to the fact that no network, like the Internet, has ever been as widespread and accessible for all. Second, platforms offer their services to users without asking for any fees to the network operator, unlike what systematically happens with other network infrastructures, such as the ones for electricity, gas, or water.

Additionally, the dimension of data collection, their management, and use, being unprecedented compared to the past, has increased the big tech companies' ability to collect and aggregate any possible data or information. This data, being then aggregated, has led to the identification of ideal types, based on recurring characteristics. Indeed, algorithms can reveal relationships and recurrences between choices, behaviors, actions, and tastes, and thus can be used to build predictive models of supply and demand for products, services, and content of various kinds. Therefore, it is fundamentally a true commercial activity, in which users' personal data is considered a commodity, and the vehicle of the exchange is advertising that providers pay to the platform.

Thus, a new era has started, where big multinational companies use algorithms and AI to govern vast communities of people. And in this context, data processed by those platforms fuels the engine of the AI-driven society[75,76], where a new kind of mass-surveillance becomes possible. Indeed, the development of AI technologies is becoming ubiquitous,

[75] From this point of view, the AI-driven society is a distinctive evolution of the 'information society'.

[76] J. M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, University of California, Davis Law Review, 2017, p. 51; A. M. Walorska, *The Algorithmic Society*, in (eds.) D. Feldner, *Redesigning Organizations Concepts for the Connected Society*, Springer, 2020; G. De Gregorio, *From Constitutional Freedoms to the Power of the Platforms: Protecting Fundamental Rights Online in the Algorithmic Society*, European Journal for Legal Studies, 2018, p. 65.

omnipresent, and seemingly omnipotent[77]. In this context, due to the enormous power that these entities are gradually acquiring – and which, actually, they continue to develop –, in the very last few years in Europe (more than in the United States), proposals and regulations[78] are now multiplying for the adoption of new rules and constraints aimed at regulating and containing such a concentration of power that have few precedents.

At the same time, as also the public actor understood the potential of AI systems in the public decision processing, a strict dependency has begun between the public and these private actors in the procurement and use of technological solutions. Indeed, in order to ensure technological development, it is clear that the public sector cannot face the transformation relying solely on internal workforces. As such, it is interesting to note the attention given to public procurement in the Coordinated plan on Artificial Intelligence[79], where it is stated that:

> 'public procurement is key in public sector AI adoption. It can also help stimulate demand and offer of trustworthy and secure AI technologies in Europe. In this context, the Commission is developing an Adopt AI program to support public procurement of AI systems and help to transform public procurement processes themselves. The program aims to help Europe's public sector to use its strong collective purchasing

[77] A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33.

[78] In this direction the EU has enacted new regulations such as the Regulation (EU) 2022/1925 of the European Parliament and of the Council of 14 September 2022 on contestable and fair markets in the digital sector and amending Directives (EU) 2019/1937 and (EU) 2020/1828 (Digital Markets Act), which has the aim of contributing to the proper functioning of the internal market by laying down rules to ensure contestability and fairness for the markets in the digital sector in general, and for business users and end users of core platform services provided by gatekeepers in particular. Moreover, the Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market For Digital Services and amending Directive 2000/31/EC (Digital Services Act) has been enacted with the aim of contributing to the proper functioning of the internal market for intermediary services by setting out harmonised rules for a safe, predictable and trusted online environment that facilitates innovation and in which fundamental rights enshrined in the European Union Charter of Fundamental Rights, including the principle of consumer protection, are effectively protected.

[79] European Commission, *Coordinated plan on artificial intelligence 2021 review*, *Annexes to the Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Fostering a European approach to Artificial Intelligence*, (April 21st, 2021).

power to act as a catalyst and stimulate demand for trustworthy AI. The public sector can lead the way in developing, purchasing, and deploying taking in use trustworthy and human-centric AI applications, for example, by utilizing public procurement of innovative solutions or by steering the development of new solutions towards its needs through pre-commercial procurement practices'.

The attention brought by the EU is also based on the consideration that governments depend on strong relationships with the private sector, a trend that is likely to intensify as technology giants widen the scope of the AI business opportunity, and governments continue to rely on external technological know-how. As such, public-private collaboration appears consistently as an essential characteristic of public technological development. This dynamic fits within a larger shift towards networked, composite collaboration between the public and the private sector in governance. Of course, private companies have been entering into legal relationships with public bodies for a long time, both for infrastructure and public service provision. But more recently, this cooperation entails the use of sophisticated systems that are directly involved in the public decision-making processes – as these systems can assume themselves the role of decision makers. Although these agreements do not involve a full transfer of responsibilities from public bodies to the private sector but rather a public-private collaboration with shared responsibilities, they still pose several risks.

Firstly, the involvement of private actors in public services has led to delayed implementation, and poor design and management choices that failed to deliver public benefits. Different studies[80] report that public-private partnerships, especially in long-term and complex projects, can lead to increased financial costs, inappropriate risk allocation,

[80] See J. F. M. Koppenjan – B. Enserink, *Public-Private Partnerships in Urban Infrastructures: Reconciling Private Sector Participation and Sustainability*, Public Administration Review, 2009, p. 284; European Court of Auditors, *Public Private Partnerships in the EU: Widespread shortcomings and limited benefits*, 2018.

misallocation of resources, over-engineered products, and under provision of citizens' needs[81]. Indeed, for all the potential benefits of public-private collaborations, the relationship is complicated from the outset. The primary duty of government is to its citizens. The primary responsibility of private companies, by contrast, is to their shareholders. Thus, these priorities can clash.

Moreover, also the delegation of public functions to private actors has to be taken into account. More than fifteen years ago, scholars already began to label this phenomenon the 'invisible handshake' according to which public actors would rely on private actors online to pursue their aims[82,83]. As Gutwirth and De Hert have observed 'how we perceive and understand our environments and interact with them and each other is increasingly mediated by algorithms'[84]. In other words, algorithms are not necessarily driven by the pursuit of public interests but are instead sensitive to business needs.

In addition, as the functionalities of the AI systems are very costly, they are usually developed by the private actor that covers them by business secrecy. This causes the public actor to be unable to access information related to the data used for training the algorithm nor to understand how the AI system works.

Said concerns are even more serious in light of the learning capabilities of algorithms, which – by introducing a degree of autonomy and thus unpredictability – are likely to undermine accountability and the human understanding of the decision-making process. For instance, the opacity of algorithms is seen as a possible cause of discrimination or differentiation between individuals when it comes to activities such as profiling and scoring

[81] A. Voorwinden, *The privatised city: technology and public-private partnerships in the smart city*, Law, Innovation and Technology, 13, 2, 2021, p. 439-463.
[82] J. R. Reidenberg, *States and Internet enforcement*, University of Ottawa Law & Techonology Journal, 2004, p. 213.
[83] G. De Gregorio, *From Constitutional Freedoms to the Power of the Platforms: Protecting Fundamental Rights Online in the Algorithmic Society*, European Journal for Legal Studies, 2018, p. 65.
[84] S. Gutwirth – P. De Hert, *Regulating Profiling in a Democratic Constitutional States*, in (eds.) M. Hildebrandt – S. Gutwirth, *Profiling the European Citizen*, 2006, p. 271.

for public purposes. The consolidation of these threats could result in a troubling process for democracy. Indeed, even if, at first glance, democratic states are open environments for pluralism flourishing through fundamental rights and freedoms, at the same time their stability can be undermined when those freedoms transform into new founding powers overcoming basic principles such as the respect of the rule of law and procedural guarantees. In this situation, there is no effective form of participation or representation of citizens in determining the rules governing their community. In other words, the creation of a private legal framework outside any representative mechanism is a threat to democracy due to the marginalization of citizens and their representatives from law-making and enforcement. For these reasons, it is currently of the uttermost importance to focus on the legislative remedies to solve the possible imbalances of powers[85].

In this context, how to best cooperate with private actors remains an interesting topic that requires further investigation from a legal point of view. Therefore, the strict link between private and public powers, as well as the influence of private actors in public affairs cannot be underestimated[86]. As such, the role of private actors in public decision-making processes through AI is investigated in chapter IV of this research.

4.2 Information sharing between private and public actors

Another aspect that should be considered is the relationship between the public and private actors in using AI systems thanks to the creation of mechanisms of data sharing. According to the European strategy for data[87]:

[85] O. Pollicino – G. De Gregorio, *Constitutional law in the algorithmic society*, Cambridge University Press, 2021, p. 3-24.

[86] L. Torchia, *Lo stato digitale*, Il Mulino, 2023.

[87] European Commission, *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*, (February 19th, 2020).

> ‘The value of data lies in its use and re-use. Currently there is not enough data available for innovative re-use, including for the development of artificial intelligence. The issues can be grouped according to who is the data holder and who is the data user, but also depend on the nature of data involved (*i.e.*, personal data, non-personal data, or mixed data-sets combining the two). Several of the issues concern the availability of data for the public good’.

Indeed, data is created by society and can serve to combat emergencies, such as floods and wildfires, to ensure that people can live longer and healthier lives, to improve public services, and to tackle environmental degradation and climate change, and, where necessary and proportionate, to ensure more efficient fight against crime. Data generated by the public sector as well as the value created should be available for the common good by ensuring, including through preferential access, that these data are used by researchers, other public institutions, small and medium enterprise’s (SMEs) or start-ups. Moreover, data from the private sector can also make a significant contribution to public goods.

However, there is currently not enough private sector data available for use by the public sector to improve evidence-driven policy-making and public services such as mobility management or enhancing the scope and timeliness of official statistics, and hence their relevance in the context of new societal developments. The recommendations of an Expert Group created by the Commission include the creation of national structures for Business to Government (B2G) data sharing, the development of appropriate incentives to create a data-sharing culture, and the suggestion to explore an EU regulatory

framework to govern the public sector's re-use for the public interest of privately held data[88].

In this direction, the EU has recently proposed and enacted different legislation aimed at fostering collaboration between the public and the private actor, as well as at enhancing transparency between them. Indeed, the promotion of AI-driven innovation is closely linked to the Data Governance Act (DGA)[89], the Data Act[90], the Open Data Directive[91] and other initiatives under the EU strategy for data, which are establishing trusted mechanisms and services for the re-use, sharing and pooling of data that are essential for the development of data-driven AI models of high quality.

As such, the functioning of a different mechanism of data sharing between the public and the private actor within the EU legal order are investigated in chapter IV of this research.

5. A path for further investigations

Algorithmic systems have contributed to the introduction of new paths for innovation, thus producing positive effects for society as a whole. Technology is also an opportunity for shaping a new paradigm of exercising the public function. AI can provide better systems of enforcement of legal rules or improve the performance of public services.

Nonetheless, the domain of inscrutable algorithms characterizing contemporary society challenges the protection of fundamental rights and democratic values while encouraging lawmakers to find a regulatory framework balancing risk and innovation, considering the

[88] European Commission, *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*, (February 19th, 2020).

[89] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act).

[90] Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonized rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act).

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

role and responsibilities of the public actor in the AI-driven age. The challenges raised by AI technologies are not limited to freedom of expression, privacy, and data protection. Constitutional democracies are under pressure to ensure legal certainty and predictability of automated decision-making processes which can collectively affect democratic values. Individuals are increasingly surrounded by ubiquitous systems that do not always ensure the possibility of understanding and controlling their underlying technologies[92]. The role of the private actor, as well as the strict interconnection with the public one for the provision of technologies and data are changing the way public decisions are reached and public interest pursued.

Therefore, as declared in chapter I the questions that will be analyzed in this research pertain to the change in the way the public actor makes decisions as a result of the adoption of AI systems. Particularly, the critical constitutional aspects of the public actor's algorithmisation are examined, as well as whether a deterioration in the relationship between the state and the citizen can be witnessed, as individuals are increasingly subjected to fully automated decisions that have a strong impact on the exercise of their fundamental rights. For this reason, it is necessary to focus on the European legal system in order to understand the direction the EU is taking in addressing this major new challenge. Indeed, leaving these issues without any safeguards would mean opening the way towards techno-determinism, allowing the actors who govern the automated systems to arbitrarily determine the standard of protection of rights and freedoms at a transnational level under the logics of digital capitalism. This is why it is critical to understand the role of regulation in the field of AI, where cooperative efforts between the public and private sector could lead to a balanced approach between risk and innovation.

[92] O. Pollicino – G. De Gregorio, *Constitutional law in the algorithmic society*, Cambridge University Press, 2021, p. 3-24.

Within this framework, the challenges raised by new automated technologies are likely to operate as a call to understand whether the current legislative provisions can guarantee the protection of fundamental rights (chapter III).

CHAPTER III: FUNDAMENTAL RIGHTS ISSUES WITHIN THE 'AI-DRIVEN PUBLIC ACTOR'

Contents: 1. Protection of fundamental rights within the 'AI-driven public actor': is it still possible?; 2. Assessment of the 'AI-driven public actor' from a fundamental rights perspective; 2.1 The right to privacy within the 'AI-driven public actor' paradigm; 2.2 The right to personal data protection in the 'AI-driven public actor' paradigm; 2.3 Non-discrimination in the 'AI-driven public actor' paradigm; 2.4 The right to a good administration in the 'AI-driven public actor' paradigm; 2.4.1 The duty of the public actor to be impartial; 2.4.2 The right of every person to have access to his or her file; 2.4.3 The right to receive a reasoned decision; 2.4.4 The accountability of the public actor; 2.5 The right to an effective remedy; 3. Regulation of AI: from the starting point to the landing point; 3.1 GDPR: The starting point; 3.1.1. Applicability of the definitions of the GDPR to AI; 3.1.2. Article 22: The regulation of the automated individual decision-making, including profiling; 3.1.2.1 The prohibition of automated decision-making processing; 3.1.2.2 Exceptions to the prohibition of automated decision-making processing and safeguards; 3.1.2.3 Automated decision-making and sensitive data; 3.1.2.4 The right to an ex-post explanation; 3.1.3. Reconciling AI and GDPR; 3.2 The attempt of the jurisprudence to answer to the open questions; 3.2.1 The initial orientation of the Italian administrative jurisprudence; 3.2.2 A change of direction of the Italian jurisprudence; 3.2.3 Ligue des Droits Humains Case; 3.2.3.1 Further considerations; 3.2.4 The Dutch SyRI Case; 3.2.4.1 The main questions before the Court; 3.2.4.2 Important remarks of the SyRI case; 3.2.5 Issues left open: the limit of the case in civil law orders; 3.3 The AI Act: the landing point?; 3.3.1 Scope of application of the AI Act; 3.3.2 Prohibited artificial intelligence practices and high risk systems; 3.3.2.1 Risk Management System; 3.3.2.2 Data Governance: Use of high-quality data sets; 3.3.2.3 Technical Documentation; 3.3.2.4 Record-keeping in order to ensure traceability and accountability; 3.3.2.5 Transparency and provision of information; 3.3.2.6 Human oversight; 3.3.2.7 Accuracy, robustness and cybersecurity; 3.3.2.8 Fundamental rights impact assessment; 3.3.3 Transparency obligations for providers and deployers of certain AI systems and GPAI models; 3.3.4 Remedies; 3.3.4.1 Lodging complaints directly with the relevant authorities; 3.3.4.2 Right to an explanation; 3.3.5 Applicability of the AI Act within the AI-driven public actor paradigm; 3.3.6 Left open questions in the protection of fundamental rights; 4. Conclusive remarks.

1. Protection of fundamental rights within the 'AI-driven public actor': is it still possible?

Considering the features of the AI-driven public actor as illustrated in chapter II, it is necessary to assess from a constitutional point of view the major impacts of AI use on individuals' possibility to exercise fundamental rights still effectively. More specifically, chapter III has the aim to understand whether the application of the most important individual rights can be guaranteed while the public actor is employing AI technologies in its decision-making process.

Particularly, the fast-growing use of AI in the fields of public welfare, policing, public security, and justice may pose the risk of ending in biased and erroneous decisions, boosting inequality, discrimination, unfair consequences, as well as undermining transparency, the right to a good administration and the right to have an effective remedy. These uses raise considerable concerns not only for specific policy areas, but also for society as a whole. Indeed, there is an increasing perception that humans do not have complete control over the transition to an AI-driven public actor and its automated decision-making processes. And, despite their predictive outperformance over analogue tools, algorithmic decisions are difficult to be understood and explained[93]. As such, algorithmic decisions could undermine procedural and substantive guarantees related to democracy and the rule of law[94].

[93] The inexplicability of AI can be associated to the so called 'black box problem' or to the difficulties in understanding the way AI reaches certain decisions. Particularly, without an adequate knowledge of how decisions are reached using algorithms, the criteria for trustworthiness cannot be satisfied. Opacity in machine learning is a complex and nuanced phenomenon that may admit of variation depending upon certain stakeholders, their interests, and sophistication. On this matter, see also F. Pasquale, *A Rule of Persons, Not Machines: The Limits of Legal Automation*, George Washington Law Review, 2019, p. 1-54.

[94] A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33. See also K. Shirley – S. Ranchordas – S. Van De Wetering, *AI failure, AI success, and AI power dynamics in the public sector*, 2024.

Therefore, being at stake such important principles and considering that full compliance with constitutional provisions is a prerequisite for using AI-driven technologies, in the context of the public actor decision processes, an analysis from a fundamental rights perspective is conducted. Specifically, the impact the new paradigm of the 'AI-driven public actor' has on the exercise of citizens' fundamental rights is investigated.

Thus, the aim of this chapter is to provide the normative basis and benchmarks for understanding whether AI tools are compatible with current EU fundamental rights framework. More specifically, this chapter introduces the general fundamental rights framework in the EU that governs the use of AI by the public actor (section 2), through the analysis of the provisions of the Charter of Fundamental Rights of the European Union (CFREU or Charter), including its specific guarantees regarding privacy, personal data protection, non-discrimination, the right to a good administration, and the right to an effective remedy. Additionally, considering the applicability of the European Charter of Human Rights (ECHR) to the European Union legal system, where relevant, the chapter also analyzes some of its provisions. Moreover, selected secondary EU legislation is presented (section 3). Specifically, 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) (GDPR) is examined, as it marked the beginning of a revolution that has seen Europe as the first to impose a set of requirements needed for implementing and using new technologies. As such, considering the large amount of personal data processed through AI also in the context of exercising public functions, GDPR's provisions are extremely relevant. Moreover, since the application of GDPR to AI in these last years has left open many interpretative questions, case law is also described in section 3. Indeed, in the absence of *ad hoc*

legislation, judges – as it often happens – has had to solve several issues connected to the use of AI and the exercise of competing fundamental rights. Finally, section 3 also examines the Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (AI Act), as it can be seen as the 'landing point' of the AI regulation after a lengthy approval process by Europe. Lastly, due to the lack of complete protection of fundamental rights in the current legislative framework, section 4 gives some conclusive remarks on the analysis of this chapter.

2. Assessment of the 'AI-driven public actor' from a fundamental rights perspective

The following description of the fundamental rights' legal framework applicable to the use of AI by the public actor is meant to start from a specific point of view. Particularly, it aims to identify concrete rights that would guarantee the use of AI systems in accordance with human rights, democracy, and the rule of law[95]. More specifically, sections 2.1-2.5 are focused on a selection of five fundamental rights whose citizens' exercise is impacted, and ultimately restricted, by the public actor's using AI. As such, the interactions between the use of AI and the citizens' exercise of the right to privacy, personal data protection, non-discrimination, good administration, and an effective remedy are taken into account. This kind of analysis is carried out considering that the selected rights are not absolute, so they can be subject to limitations in line with Art. 52(1) of the Charter. Accordingly, the limitation of citizens' fundamental rights is not unlawful *per se*. Indeed, as the interest of

[95] D. Leslie – C. Burr – M. Aitken – J. Cowls – M. Katell – M. Briggs, *Artificial intelligence, human rights, democracy, and the rule of law: a primer*, The Council of Europe and the Alan Touring Institute, 2021.

the public actor in making public decisions faster and more efficiently exists, the need to strike a balance with competing rights is urgent. Thus, before analyzing to what extent the fundamental rights are impacted by the use of AI and whether their limitations are compatible with Art. 52 of the Charter, the general steps that need to be followed according to the aforementioned provision are presented[96]. Particularly, pursuant to Art. 52(1) of the Charter:

> 'Any limitation on the exercise of the rights and freedoms recognized by this Charter must be provided for by law and respect the essence of those rights and freedoms. Subject to the principle of proportionality, limitations may be made only if they are necessary and genuinely meet objectives of general interest recognized by the Union or the need to protect the rights and freedoms of others'.

The interpretation of this important provision is based on the case law of the Court of Justice of the EU (CJEU) as:

> 'it is well established in the case-law of the Court that restrictions may be imposed on the exercise of fundamental rights, in particular in the context of a common organization of the market, provided that those restrictions in fact correspond to objectives of general interest[97] pursued by the Community and do not constitute, with regard to the aim pursued, disproportionate and unreasonable interference undermining the very substance of those rights'[98,99].

Accordingly, Art. 52(1) contains three different elements: (i) a procedural rule (limitations on rights 'must be provided for by law'); (ii) a rule on the justifications for limiting rights

[96] Art. 52(3) of the Charter.
[97] The reference to general interests recognized by the Union covers both the objectives mentioned in Art. 3 of the Treaty on European Union (TUE) and other interests protected by specific provisions of the Treaties such as Art. 4(1) of the TUE and Art. 35(3), 36 and 346 of the Treaty on the Functioning of the European Union (TFUE).
[98] CJUE, *Kjell Karlsson and Others,* C-292/97, (2000).
[99] S. Peers – S. Prechal, *Article 52,* in (eds.) S. Peers - T. Hervey - J. Kenner - A. Ward, *The EU Charter of Fundamental Rights: A Commentary*, Oxford, 2021, p. 1611–1674.

('objectives of general interest recognized by the Union or the need to protect the rights and freedoms of others') and a number of interlinked rules on the balancing test to be applied as between rights and limitations (the obligation to 'respect the essence of' the rights; the 'principle of proportionality'); (iii) the requirement of necessity[100].

On this merit, the CJEU has emphasized that any limitation on the exercise of the rights and freedoms recognized by in the Charter must respect 'the essence' of those rights and freedoms. This means that fundamental rights can be limited to a certain extent, but not completely disregarded. Therefore, in balancing different rights, once it has been established that the inalienable, essential core of a right is not violated by a measure, the next step is to conduct the necessity and proportionality test outlined in the Charter in respect of non-core aspects of that right. Particularly, under those tests, any interference with a Charter right can be justified if the given legitimate aim could not have been obtained by other means that interfere less with the guaranteed right.

Similar requirements are also imposed by the ECHR, as interpreted by the European Court of Human Rights (ECtHR). These include the 'essence of a right' concept, which can be derived from the object and purpose of the ECHR as a whole. Following this theory, the same Court, in respect to the use of new technologies, observed in *S. and Marper v. the UK*[101] that States should 'strike a right balance' between protecting fundamental rights and developing new technologies.

Therefore, given the wide range of applications of AI systems by the public actor as presented in chapter II and the possibility to restrict fundamental rights under a strict necessary and proportionality test, the following sections assess whether a balance exists between the public actor's interests in using AI and selected fundamental rights, also in order to preserve and respect 'the essence' of those rights and freedoms.

---

[100] *Ibid*.

[101] ECtHR, *S. and Marper v. The United Kingdom*, App. 30562/04 and 30566/04, (2008).

2.1 The right to privacy within the 'AI-driven public actor' paradigm

Art. 7 of the Charter provides for the right to respect for private life as it states that:

> 'Everyone has the right to respect for his or her private and family life, home and communications'.

The analogous Art. 8 of the ECHR specifies that:

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

According to the wording of Art. 7(1) of the Charter (as well as of Art. 8(1)), the concept of 'private life' or 'privacy' is complex and broad, and not susceptible to an exhaustive definition. It covers the physical and psychological integrity of a person, and can, therefore, embrace multiple aspects of the person's physical and social identity. These aspects, rather being guaranteed just in the innermost private sphere of the individual, can be protected also in public spaces. Indeed, the concept of privacy can cover a zone of interaction between a person and others even in a public context. For instance, the ECtHR has used the concept of 'reasonable expectation of privacy' – referring to the extent to which people can expect privacy in public spaces without being subjected to surveillance – as one of the factors, albeit not necessarily a conclusive one, to decide on a violation of the right to respect for private life[102].

[102] However, not being an absolute right, its relevance and scope of application, appears to be limited. For instance, the mere fact that participants in assemblies are out in public does not mean that their privacy

However, as the Explanatory Notes to the CFREU[103] articulate, since Art. 7 of CFREU (like Art. 8 of ECHR) is a qualified right, any interference with the protected rights is prohibited unless it falls within the limitations permitted by Art. 52 of CFREU[104].

Following this path, in relation to the interest of the public actor and the citizens' reasonable expectation of privacy, Art. 8(2) of the ECHR points out the elements necessary for striking a balance between them stating that the right to privacy can be restricted just (i) in accordance with the law and (ii) if it is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country, for the prevention of disorder or crime, for the protection of health or morals, or for the protection of the rights and freedoms of others.

Focusing on the debate about the interference of privacy when using AI systems (specifically by the public actor), this right is one of the core fundamental, as it strives to protect the autonomy and human dignity of individuals by granting them a personal sphere in which they can freely develop their personalities, think, and shape their opinions. Moreover, since the ability of AI to make predictions about a person's behavior, state of mind, and identity, the right to privacy can be invoked to avoid the unreasonable intrusion to individuals' private sphere by external actor, as well as the altering of individuals' behavior upon suspicion that he/she is being observed or analyzed[105].

The extent of these risks can be well understood through an example. For instance, the use by the public actor of live facial recognition technology that involves the biometric processing of facial images taken in a public place for the purpose of determining constantly a person's identity (one-to-many identification), in absence of a strict

cannot be infringed. The same applies to the monitoring of social media to glean information about participation in peaceful assemblies.

[103] Explanations Relating to the Charter of Fundamental Rights, (2007/C 303/02).

[104] J. Vested-Hansen, *Article 7 (Private Life, Home and Communications)*, in (eds.) S. Peers – T. Hervey – J. Kenner – A. Ward, *The EU Charter of Fundamental Rights: A Commentary*, Oxford, 2021, p. 151–194.

[105] D. Leslie – C. Burr – M. Aitken – J. Cowls – M. Katell – M. Briggs, *Artificial intelligence, human rights, democracy, and the rule of law: a primer*, The Council of Europe and the Alan Touring Institute, 2021.

proportionality test, as required by Art. 8(2) of the ECHR, constitutes an interference with the right to respect for private life of citizens[106].

Thus, within the new paradigm of the AI-driven public actor, the right to privacy should serve as a safeguard against arbitrary and continuous state surveillance and control. Particularly, the guarantee of the right to privacy should prevent governments from carrying out unwanted monitoring, collecting data, or invading people's private lives. To avoid excessive government control and to advance democratic norms, privacy aids in maintaining a balance between security concerns and the preservation of civil freedoms. Ultimately, it protects citizens against tyrannical governments, advances the rights to free speech and association, and finds solutions to legal concerns[107] brought on by the AI-driven public actor.

2.2 The right to personal data protection in the 'AI-driven public actor' paradigm

In the paradigm of the AI-driven public actor, the right to personal data protection, as guaranteed by Art. 8 of the Charter, constitutes an important safeguard to citizens' liberties and freedoms. Indeed, considering that the public actor has the privilege to acquire *ex lege* a wide availability of personal data and information, the risk that it may use such information to practice some form of mass surveillance is high. Particularly, the increasing presence of big databases has been the perfect companion to and one of the main facilitating factors behind the current build-up of AI. From recidivism scoring, social funds allocation, and CV screening, to FRT, there is a myriad of AI applications that are fueled

[106] European Union Agency for Fundamental Rights, *Facial recognition technology: fundamental rights considerations in the context of law enforcement* (November 2019).
[107] S. Gilani – A. Al-Matrooshi – M. Khan, *Right of Privacy and the Growing Scope of Artificial Intelligence. Current Trends*, Law and Society, 2023, p. 1-11.

by personal data[108]. Moreover, the processing of such information through AI systems allows the public actor to determine the terms and conditions through which citizens are granted access to certain public services. Particularly, the most advanced systems of AI based on the analysis of data are variously and widely used to support public decision-making in order to, for instance, facilitate automated control of social funds, infer insights from large datasets in the health sector, preventing and detecting school absence and early leaving.

Therefore, in order to provide specific guidelines in the use of AI by the public actor, it is essential to analyze the content of Art. 8 of the Charter, which states that:

> 'Everyone has the right to the protection of personal data concerning him or her.
>
> Such data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law.
>
> Everyone has the right of access to data which has been collected concerning him or her, and the right to have it rectified'.

Particularly, section 1 is a mandate to the legislator to regulate the processing of personal data so that the interests of citizens are adequately protected. In order to do that, section 2 points out specific requirements (*i.e.*, data must be processed fairly for specified purposes and on the basis of the consent of the person concerned or some other legitimate basis laid down by law), which much be taken into consideration in the legislation. Thus, Art. 8 establishes a duty upon the legislator to enact and structure data protection regulations in a way that they are compatible with the principles enshrined in section 2[109].

---

[108] N. Menéndez González, *The Rights to Privacy and Data Protection and Facial Recognition Technology in the Global North,* in (eds.) A. Quintavalla – J. Temperman, *Artificial Intelligence and Human Rights*, Oxford University Press, 2023, p. 136.

[109] N. Marsch, A*rtificial Intelligence and the Fundamental Right to Data Protection: Opening the Door for Technological Innovation and Innovative Protection,* in (eds.) T. Wischmeyer – T. Rademacher, *Regulating Artificial Intelligence*, Berlin, 2020.

In this framework, it seems that AI comes into conflict with the basic core elements of the data protection right. For instance, considering that the use of the current conventional machine-learning-based AI is based on large data volumes that have been regularly collected for other purposes, there is a clash with the requirements under Art. 8(2) of the Charter, as it would be difficult to ensure that the data subject is always informed of the specified purposes of the processing, as well as he or she has given its own consent. Moreover, the idea of regulating every data processing step is diametrically opposed to the so-called black box problem[110]. Additionally, the concern relating to the loss of control reflected in the data protection law faces the partially uncontrolled and uncontrollable form of AI. Furthermore, even if the public actor could lay down by law other legitimate basis for the processing of a large amount of personal data, there is a risk that it would abuse its position by making legitimate processing of personal data that, because of its extent, could lead to mass surveillance. Therefore, it becomes clear that there is at least a certain tension between the use of AI by the public actor and the personal data protection right.

Thus, in view of the risks and dangers associated with the use of AI, it may be disputed even the possibility, from a fundamental rights perspective, to consider the use of AI by the public actor in violation of Art. 8 of the ECHR. However, this assessment cannot be limited to the analysis of the constitutional principles[111]. Therefore, in the field of data protection, it is necessary to refer to secondary legislation, and particularly to the GDPR, as it is examined in section 3.

---

[110] F. Pasquale, *The black box society: the secret algorithms that control money and information*, Harvard University Press, 2015. For purposes of this chapter, black box AI refers to any natural language processing, machine learning, textual analysis, or similar software which uses data not accessible to the data subject, and/or which deploys algorithms which are either similarly inaccessible, or so complex that they cannot be reduced to a series of rules and rule applications comprehensible to the data subject.

[111] N. Marsch, A*rtificial Intelligence and the Fundamental Right to Data Protection: Opening the Door for Technological Innovation and Innovative Protection,* in (eds.) T. Wischmeyer – T. Rademacher, *Regulating Artificial Intelligence*, Berlin, 2020.

2.3 Non-discrimination in the 'AI-driven public actor' paradigm

European non-discrimination fundamental right is essential for safeguarding citizens' freedoms and liberties in the context of the use of AI by the public actor. Art. 2 of the TUE provides that non-discrimination is one of the fundamental values of the EU, and Art. 10 of the TFEU requires the Union to combat discrimination on a number of grounds. Moreover, Art. 21 of the Charter provides for non-discrimination[112], stating that:

> 'Any discrimination based on any ground such as sex, race, color, ethnic or social origin, genetic features, language, religion or belief, political or any other opinion, membership of a national minority, property, birth, disability, age or sexual orientation shall be prohibited.
>
> Within the scope of application of the Treaty establishing the European Community and of the Treaty on European Union, and without prejudice to the special provisions of those Treaties, any discrimination on grounds of nationality shall be prohibited'.

The very two-limb structure of Art. 21 of the Charter, whereby nationality discrimination is treated separately, and differently, from status discrimination, signals the distinctive contours that nationality discrimination has been given in the EU. On the one hand, it has been central for developing very extensive protection for EU citizens, although in order to fall within the scope of EU law there has normally been a need for a cross-border dimension before it is engaged and the CJEU has moved to further curtail equal treatment protection for non-economically active, albeit lawfully residing, EU citizens[113]. On the other hand, throughout the development of EU citizenship rights, a restrictive approach has been taken by the CJEU by denying its protection to third-country nationals.

[112] European Union Agency for fundamental rights, *Getting the Future Right Artificial Intelligence and Fundamental Rights*, Luxembourg: Publications Office of the European Union, (December 2020).
[113] CJUE, *Dano*, C-333/13, (2014) and CJUE, *Alimanovic*, C-67/14, (2015).

Discrimination has a well-established, albeit evolving, EU law grammar, with distinctive configurations in relation to particular grounds. Its core components consist of the concepts of direct[114] and indirect discrimination[115]. In the classical structure, direct discrimination can usually not be justified or is subject to stricter and often legislatively defined justification as compared to indirect discrimination, where a broader range of 'objective justifications' are generally permitted. Consideration is also given to departures from symmetrical equal treatment in relation to the disadvantaged group. This may range from requiring only the disadvantaged group to be covered by discrimination protection (*e.g.*, disability in EU law) to permit certain kinds of positive action as a tightly construed exception to an otherwise symmetrical application of the status discrimination prohibition. Other features include protecting complainants through the special burden of proof regimes and victimization protection as well as protection for special equality agencies and constructing harassment as a specific form of discrimination[116].

In this framework, such broad protection afforded by Art. 21 seems difficult to be ensured in the use of AI by the public actor. Indeed, AI systems are capable of reproducing and augmenting the patterns of discriminatory treatment that exist in the society in which they are created and used. This can occur when the stereotyping biases and blind spots of system developers shape the choices made in the design and deployment of systems. It can also occur when historical structures of inequality and discrimination become entrenched in the datasets that are used to train AI and machine learning models. Where AI – and consequently the public actor – relies on such biased information, discriminatory

[114] Direct discrimination shall be taken to occur where one person is treated less favorably than another is, has been or would be treated in a comparable situation, on any of the prohibited grounds.
[115] Indirect discrimination shall be taken to occur where an apparently neutral provision, criterion or practice would put persons of a prohibited status at a particular disadvantage compared with other persons, unless that provision, criterion or practice is objectively justified by a legitimate aim and the means of achieving that aim are appropriate and necessary.
[116] H. Eklund, *Article 21*, in (eds.) S. Peers – T. Hervey – J. Kenner – A. Ward, *The EU Charter of Fundamental Rights: A Commentary*, Oxford, 2021, p. 613–638.

human decisions that produced a dataset can lead to discriminatory algorithmic decisions and behaviors[117]. More specifically, direct or indirect discrimination through the use of algorithms that involve big data is considered as one of the most pressing challenges in the use of AI-driven technologies. Bias and discrimination, including gender-based discrimination, in data-supported algorithmic decision making can occur for several reasons and at many levels in the AI systems. They are difficult to detect and mitigate. Often, the quality of the data and biases within it are the source of potential discrimination and unfair treatment. The discriminatory effects generated on certain groups are, in practice, very difficult for individuals to challenge. So far, only a limited number of court cases have dealt with discrimination relating to AI systems[118]. Moreover, studies have

[117] D. Leslie – C. Burr – M. Aitken – J. Cowls – M. Katell – M. Briggs, *Artificial intelligence, human rights, democracy, and the rule of law: a primer*, The Council of Europe and the Alan Touring Institute, 2021.

[118] In relation to case law about the use of AI by the public actor, it should be recalled the decision of the District Court of the Hague, 6 March 2020, n. 865, in which, the Court of The Hague rendered its judgment in *NCJM et al. and FNV v. The State of the Netherlands*. The case challenged the Dutch government's use of System Risk Indication (SyRI) – an algorithm designed to identify potential social welfare fraud. The Court ruled that neither the legislation governing SyRI nor its use met the requirements laid down in Art. 8(2) of the ECHR for an interference with the exercise of the right to private life to be necessary and proportionate. Moreover, the Court found that the algorithms used discriminatory patterns. This is one of the first judgments in the world addressing the human rights implications of the use of AI in the public sector and states' respective obligations to ensure transparency of AI processes (for further analysis please see section 3.2.3 of this chapter).
Moreover, even if it is a US case, *State v. Loomis*, 881 N.W. 2d 749 (2016) should be recalled. In early 2013, Wisconsin charged Eric Loomis with five criminal counts related to a drive-by shooting in La Crosse. Loomis denied participating in the shooting, but he admitted that he had driven the same car involved later that evening. Loomis pleaded guilty to two of the less severe charges. In preparation for sentencing, a Wisconsin Department of Corrections officer produced a PSI that included a COMPAS risk assessment. COMPAS assessments estimate the risk of recidivism based on both an interview with the offender and information from the offender's criminal history. As the methodology behind COMPAS is a trade secret, only the estimates of recidivism risk are reported to the court. At Loomis's sentencing hearing, the trial court referred to the COMPAS assessment in its sentencing determination and, based in part on this assessment, sentenced Loomis to six years of imprisonment and five years of extended supervision. Loomis filed a motion for post-conviction relief in the trial court, arguing that the court's reliance on COMPAS violated his due process rights. Because COMPAS reports provide data relevant only to particular groups and because the methodology used to make the reports is a trade secret, Loomis asserted that the court's use of the COMPAS assessment infringed on both his right to an individualized sentence and his right to be sentenced on accurate information. Loomis additionally argued on due process grounds that the court unconstitutionally considered gender at sentencing by relying on a risk assessment that took gender into account.17 The trial court denied the post-conviction motion, and the Wisconsin Court of Appeals certified the appeal to the Wisconsin Supreme Court. The Wisconsin Supreme Court affirmed. Writing for the court, Justice Ann Walsh Bradley rejected Loomis's due process arguments. Justice Bradley found that the use of gender as a factor in the risk assessment served the nondiscriminatory purpose of promoting accuracy and that Loomis had not provided sufficient evidence that the sentencing court had actually considered gender. Moreover, as COMPAS uses only publicly available data and data provided by the defendant, the court concluded that Loomis could have denied or explained any information that went into making the report and

highlighted the potential for discrimination prompted by the use of AI-systems across the public sector[119].

In response to these considerations and concerns, the public actor should adopt a precautionary approach in the adoption and regulation of AI that balances the realization of the opportunities presented by AI while ensuring that risks to human beings and human interests are minimized to the extent possible. In contexts where there is uncertainty about the level or impact of potential risks, governments should apply a higher degree of regulatory oversight and monitoring of AI systems and be prepared to prohibit its use[120].

2.4 The right to a good administration in the 'AI-driven public actor' paradigm

In analyzing the set of fundamental rights that have to be guaranteed by the public actor while using AI for automated decision processing, also the right to a good administration has to be considered. According to Art. 41 of the Charter:

> 'Every person has the right to have his or her affairs handled impartially, fairly and within a reasonable time by the institutions, bodies, offices and agencies of the Union.
>
> This right includes:
>
> (a) the right of every person to be heard, before any individual measure which would affect him or her adversely is taken;
>
> (b) the right of every person to have access to his or her file, while respecting the legitimate interests of confidentiality and of professional and business secrecy;
>
> (c) the obligation of the administration to give reasons for its decisions.

---

therefore could have verified the accuracy of the information used in sentencing. However, despite the final judgment, the case is emblematic in assessing for the first time the concerns related to the application of non-discrimination right in the case a decision is taken (or even suggested) by an AI technology.

[119] European Union Agency for fundamental rights, *Getting the Future Right Artificial Intelligence and Fundamental Rights*, Luxembourg: Publications Office of the European Union, (December 2020).

[120] D. Leslie – C. Burr – M. Aitken – J. Cowls – M. Katell – M. Briggs, *Artificial intelligence, human rights, democracy, and the rule of law: a primer*, The Council of Europe and the Alan Touring Institute, 2021.

Every person has the right to have the Union make good any damage caused by its institutions or by its servants in the performance of their duties, in accordance with the general principles common to the laws of the Member States.

Every person may write to the institutions of the Union in one of the languages of the Treaties and must have an answer in the same language'.

Thus, Art. 41 guarantees basic procedural rights[121] which should protect individuals' rights and interests in situations where institutions, bodies and agencies of the EU take administrative actions against individuals[122].

Accordingly, Art. 41(1) provides that every person has the right to have his or her affairs handled impartially, fairly and within a reasonable time by the institutions, bodies, offices, and agencies of the Union. The CJEU has held that the requirement of impartiality in Art. 41(1) covers subjective impartiality, which precludes bias or personal prejudice, and objective impartiality, under which there must be sufficient guarantees to exclude any legitimate doubt as to possible bias on the part of the institution concerned[123,124].

Moreover, this right includes the right of every person to be heard (Art. 41(2)(a)), the right of an individual to have access to their file (Art. 41(2)(b)) and the obligation of any public authority to give sufficient reasons for its decisions (Art.41(2)(c)). Access to the file facilitates understanding of the evidentiary basis on which a decision has been made, and/or of the reasons underlying it. This places the individual in a better position to put forward counterarguments when exercising the right to be heard and the right to an effective remedy. Therefore, the obligation to give reasons makes, from the perspective

[121] A. Jackiewicz, *Prawo do dobrej administracji jako standard europejski (The Right to Good Administration As European Standard)*, Toruń, 2008, p. 58.
[122] M. Lais, *Das Recht auf eine gute Verwaltung unter besonderer Berücksichtigung der Rechtsprechung des Europäischen Gerichtshofs*, Zeitschrift für Europarechtliche Studien, 2002.
[123] CJEU, *Ziegler v. Commission*, C-439/11, (2013); CJEU, *Spain v. Commission,* C-521/15, (2017); CJEU, *Icap v. Commission*, T-180/15, (2017).
[124] P. Craig, *Article 41*, in (eds.) S. Peers – T. Hervey – J. Kenner – A. Ward, *The EU Charter of Fundamental Rights: A Commentary*, Oxford, 2021.

of the individuals affected, the decision-making process more transparent, so that the individual can understand why a measure or action has been taken. Transparency is also an enabling principle that provides foundations for other rights, including the exercise of the right to an effective remedy[125].

Lastly, Art. 41(3) establishes the principle of accountability of public administration as it states that every person has the right to have the Union make good any damage caused by its institutions or by its servants in the performance of their duties. In this context, in order to understand whether the right to a good administration can be still guaranteed in the new paradigm of the AI-driven public actor, the following paragraphs delve into these four elements of Art. 41, and specifically into the duty of the public administration to be impartial (section 2.5.1), the right of every person to have access to his or her file (section 2.5.2) the right of the individuals to receive a reasoned decision (section 2.5.3) and the principle of accountability (section 2.5.4)[126].

2.4.1 The duty of the public actor to be impartial

According to Art. 41(1) of the Charter every person has the right to have his or her affairs handled impartially, fairly and within reasonable time by the institutions, bodies, offices, and agencies of the EU. These are very crucial features of the administrative procedure as they influence relations between individuals and administrations, as well as build individuals' confidence in the public actor[127].

[125] European Union Agency for fundamental rights, *Getting the Future Right Artificial Intelligence and Fundamental Rights*, Luxembourg: Publications Office of the European Union, (December 2020).

[126] Note that only the relevant rights and duties for the use of AI by the public actor have been selected. Therefore, the following analysis cannot be considered as a comprehensive description of Art. 41 of the Charter.

[127] M. Szydło, *Prawo do dobrej administracji jako prawo podstawowe w unijnym porządku prawnym (The Right to Good Administration as Fundamental Right in the Union's Legal Order)*, Studia Europejskie, 2004, p. 95.

Accordingly, impartiality means that actions are free of pressure, partiality or interest and implies 'absence of discrimination'[128], as well as requires that administration does not act arbitrarily. Moreover, it means that there should be no doubts about the objectivity of officials who take part in decision-making processes. Their actions cannot be determined by personal, family, financial and national interest and must be free of political influence. Otherwise, the officials cannot take part in further proceedings and must be excluded[129]. Therefore, proceedings as well as the final decisions should be fair. As such, the principle implies that officials must take into account all factual and legal information available at the time of the decision-making process[130].

In the context of the AI-driven public actor, it may seem that replacing human decision-makers with AI could imply better enforcement of the principle of impartiality. Indeed, AI should be able to filter out the relevant factors to take an administrative decision by itself, therefore, eliminating the biases inherent to any kind of human decision-making. Nonetheless, a blind belief in the inherent impartiality of algorithms is misguided, as the ability to reason logically doesn't preclude discriminations[131,132]. Particularly, the use of automated decision-making by the public actor poses a further challenge to the idea that all individuals irrespective of their status must have equal access to rights, and that in accessing these rights 'like cases be treated alike'. It cannot be assumed that automated systems will discriminate in ways similar to humans, or familiar known patterns of discrimination. AI and automated systems are valued precisely because of their ability to process data at scale and find unintuitive connections and patterns between people.

[128] J. Ponce, *Good Administration and Administrative Procedures*, Indiana Journal of Global Legal Studies, 2005, p. 567.
[129] G. Krawiec, *Europejskie prawo administracyjne (The European Administrative Law)*, Warszawa, 2009, p. 108.
[130] CJUE, *Estabelecimentos Isidoro M. Oliveira SA v. Commission of the EC*, T-73/95, (1997).
[131] On the discriminatory features of AI systems, please see also section 2.3 of this chapter.
[132] L. Van Wichelen - J. C. Devogelaere, *Artificial Intelligence: From public discrimination to public administration*, The EUTOPIA Student Think Tank (EUSTT), 2023.

However, automated decision-making systems can undermine the principle of impartiality in different ways as they may: i) explicitly incorporate and rely on various static factors and/or immutable characteristics, such as socio-economic status, employment and education, postal codes, age, or gender; or ii) take such matters into account indirectly, for example by learning the relevance of variables correlations[133]. Therefore, when designing an algorithm, making it as fair as possible will be one of the main concerns of the developers. Both code-driven and data-driven algorithmic systems have very different problems with adhering to the principle of impartiality. The major problems, however, stand or fall with the quality of the input data. If the data contains certain discriminatory relationships, the algorithm will also be biased, so there is a need to build some form of fairness into the algorithm. Simply leaving out data linked to factors that lead to discrimination of certain groups, such as gender, race, or religious conviction, does not always lead to this discrimination disappearing. In addition to possible discriminatory input data, there is also the problem of discriminatory feature selection: one must instruct the algorithm on which factors have to be considered in order to reach a decision. Indeed, even if every dataset is factually unbiased, drawing correlations can mean discrimination. As such, it will probably be impossible, both for human officials and for algorithms, to eliminate discrimination[134]. Therefore, even if it would be possible to not incorporate inappropriate variables in the automated system, it should be still considered that automation can infer rules from historical patterns and correlations. As such, even when variables, such as race, are not used in the learning process, a machine can still produce racially or otherwise biased assessments.

[133] M. Zalnieriute – L. Bennett Moses – G. Williams*, The Rule of Law and Automation of Government Decision-Making*, Modern Law Review, 2019, p. 245.
[134] L. Van Wichelen – J. C. Devogelaere, *Artificial Intelligence: From public discrimination to public administration*, The EUTOPIA Student Think Tank (EUSTT), 2023.

As these examples demonstrate, algorithms are far from neutral pieces of technology. Consequently, in understanding the benefits and challenges of the public automated decision processing, it is crucial to consider both the context of the decision, and the type of system deployed. A system with pre-programmed rules can ensure that decisions are based on factors recognized as legally relevant and hence avoid or minimize the risk of non-impartial decisions. However, procedural rights and opportunities to check and rectify data on which the decision relies are crucial, as it is ensured that the logic of the system accurately reflects the law. Therefore, the challenges posed by systems based on rules inferred from data are different. Here, the role of humans is limited to setting parameters, selecting data (possibly biased due to flawed human collection practices), and deciding which variables to use as a basis for analysis. Unless the humans involved in these processes have a deep understanding of the legal context in which a decision is made, systems may fail in practice to meet the standard of impartiality, as set in Art. 41 of the Charter[135].

Considering the above-mentioned criticalities, additional regulation should be considered to guarantee fundamental rights and fairness in the algorithmic decision-making processing.

2.4.2 The right of every person to have access to his or her file

Under Art. 41(2)(b) of the Charter the right of every person to have access to his or her file is established. Access to the file may be relevant before the decision is made by the administration, or after it has been made when an applicant seeks to challenge the decision by judicial review. Access facilitates understanding of the evidentiary basis on which the decision is to be made or has been made, and of the reasoning underlying it,

[135] M. Zalnieriute – L. B. Moses – G. Williams, *The Rule of Law and Automation of Government Decision-Making*, Modern Law Review, 2019, p. 245.

thereby placing the individual in a better position to put counterarguments when exercising the right to be heard or challenging the decision by way of judicial review. Access to the file and access to documents as protected by Art. 15(3) TFEU and Art. 42 of the Charter can function as alternate routes to the same goal. EU law accords access to the file as part of the rights of the defense.

However, the general principle of access to the file is subject to a number of limitations. There is no access to business secrets and confidential information, but the public actor cannot make a general reference to confidentiality to justify a total refusal to disclose documents in its file, nor can it give blank pages on the ground that they contained business secrets without providing a more comprehensible non-confidential version, or a summary of the documents. There is no general principle that the parties must receive copies of all documents taken into account in the case of other people. Moreover, there is no right to access documentation that is irrelevant and bears no relation to the allegations of fact or law in the statement of objections. Lastly, the right of access to the file cannot be used to circumvent the state aid rules concerning challenges to such decisions raised by third parties[136].

In this perspective, the right of access to information can be outlined as an instrument against discretional exercise of government power, which is recognized to all individuals in public interest. Therefore, the exercise of the right to access file and information is strictly liked with the duty of transparency of the public actor, and, as such, it should underpin the government use of AI[137].

[136] P. Craig, *Article 41,* in (eds.) S. Peers – T. Hervey – J. Kenner – A. Ward, *The EU Charter of Fundamental Rights: A Commentary*, Oxford, 2021, p. 1125–1152.

[137] According to the European Parliament, the principle of transparency means that 'it should always be possible to supply the rationale behind any decision taken with the aid of AI that can have a substantive impact on one or more persons' lives'.

However, in the context of AI, the duty of transparency can clash with the opacity of the technology. Consequently, although algorithms may make the functioning of the public actor more efficient, they can also make it less transparent[138,139].

Specifically, the opacity of algorithms may come down to technical, legal, or organizational reasons. First, it may be because of technical reasons. Particularly, it may stem from the complexity of the algorithms or their dynamic nature. With certain algorithms, it is difficult to know precisely how they work or what data they take into account in producing a particular result. Understanding them may require devoting many resources and having particular knowledge or specific skills. Moreover, due to their dynamic nature, even if it would be possible to discover how an algorithm is designed, this would be based on its characteristics in a particular moment and may have subsequently changed. Additionally, from a legal standpoint, the opacity of the algorithms may be strengthened due to the existence of contractual clauses or regulations that limit access to information to protect other assets or rights, such as trade secrets, intellectual and industrial property, personal data protection and public safety[140]. It may also be enhanced to protect the confidentiality or secrecy of the decision-making process to avoid revealing decisions still in the making. Moreover, it should be considered that the algorithms used are often not programmed by the public sector but by the private one who wants to protect them, in order to avoid losing a competitive advantage. That is why it is common for contractors to declare their

---

138 C. Ramió Matas, *Inteligencia artificial y administración pública. Robots y humanos compartiendo el servicio público,* Catarata, 2019.

139 S. Ranchordas, *Experimental regulations for AI: sandboxes for morals and mores*, Morals & Machines 1.1, 2021, p. 86-100.

140 In this context, it is interesting to recall the recent case CJEU, *Schufa*, C-634/21, (2023), where the request has been made after a proceeding between OQ and the Land Hessen (Federal State of Hesse, Germany) concerning the refusal of the *Hessischer Beauftragter für Datenschutz und Informationsfreiheit* (Data Protection and Freedom of Information Commissioner for the Federal State of Hesse, Germany; 'the HBDI') to order SCHUFA Holding AG to grant an application lodged by OQ seeking to access and erase personal data concerning her. That refusal was actually based also on Schufa's intellectual property rights. However, in the judgement brought in front of the CJUE, the Court held that right of transparency, as well as the other rights provided for by the GDPR had to be considered predominant.

algorithm codes confidential. Lastly, from an organizational perspective, opacity may be the result of a total lack of information about algorithms. In some cases, this may be because the public authorities that use the algorithms do not have the source code. In other cases, it may be because the public actor has not formalized its decision to use a certain algorithm, the purpose behind them, the sources or the results obtained through the algorithms.

Ultimately, the opacity of algorithms can have dire consequences for citizens' exercise of fundamental rights. On the one hand, the government may not know how the algorithms work and, therefore, how the decisions they adopt are reached. This may have consequences regarding compliance with the principles and regulations governing public actor's activities, as well as in terms of monitoring the administrative action and the ways public decisions are taken. On the other hand, opacity can also lead to discrimination arising from biases that the algorithms may contain that cannot be detected by public employees. While algorithms can help public employees avoid cognitive biases that may unconsciously form part of their decisions, the opacity of algorithms may make this attitude impossible to detect. This may also hinder the identification of errors generated by algorithms to the extent that, eventually, any decision taken by or with the intervention of the AI is considered correct owing to the belief that a computer works better than a human mind[141]. Furthermore, the use of algorithms that are designed by third parties creates the risk that the companies that develop them and have access to data end up taking decisions on behalf of the government.

Considering all these issues, in order to avoid the opacity of algorithms, the public actor must adopt certain measures to aid transparency. These measures should allow not only

[141] R. Brauneis – E. P. Goodman, *Algorithmic transparency for the smart city*, Yale Journal of Law & Technology, 2019, p. 104-176.

access to algorithms, but also knowledge of their content and an understanding of how they work and the basis for any decisions taken using them.

2.4.3 The right to receive a reasoned decision

Under Art. 41(2)(c) of the Charter, the right to give reasons for decisions is another component of the right to good administration.

The duty to give reasons has the functions of giving the courts the possibility of exercising their power to review the legality of the decision, and the individuals the opportunity to protect their rights (receiving enough information to be able to determine whether the decision is well-founded). In this regard, the CJEU often invokes Art. 47 of the CFREU to support the requirement of reason-giving. Particularly, according to Art. 47 of the Charter sufficient reasoning is both a procedural and a substantive guarantee which allows the plaintiffs to decide on the potential success of their claim and thus to effectively prepare a defense. Accordingly, reasoning 'must enable a person to defend his or her rights in the best possible conditions and to decide, with full knowledge of the relevant facts, whether there is any point in applying to the court with jurisdiction'[142]. In providing the reasoning, the authority is required to explain the facts and legal considerations having decisive importance in the context of their decisions.

However, the reasoning does not have to include all points of fact and law, since it has to be assessed considering its context and all the legal rules governing the matter in question.

According to the CJEU, the duty to state reasons varies considering the type of act, the nature[143], the content of the decision, the interests of the individuals affected by such

[142] CJEU, *R.N.N.S. and K.A. v. Minister van Buitenlandse Zaken*, C-225/19 and C-226/19, (2020).
[143] CJEU, *W. Beus GmbH & Co. v. Hauptzollamt München*, C-5-67 (1968).

decision, the specific context[144], and the legal rules[145] governing the matter in question. For instance, in the case of acts of general application, the reasoning has to provide the legal justification, an explanation of the situation that led to the adoption of the act and the objectives which it is intended to achieve. A more stringent duty to state reasons is required for acts addressed to individuals. Indeed, the addressee must be able to assess the lawfulness of the act affecting him or her and possibly challenge it. Another element affecting the statement of reasons is the discretion of the body adopting the act. In the case of discretionary acts, the reasoning should be more rigorous, to facilitate the understanding of the reasons, as well as the exercise of judicial review, and may not be limited to an indication of the factors analyzed. Moreover, in case of discretionary acts it is necessary to indicate objective and predetermined criteria on which such acts can be adopted. Also, according to the CJEU[146], when the act is part of an established case law, measures may be reasoned in a summary manner (as long as the judicial review is possible). Conversely, if the body states a new principle or applies it in a different way (*i.e.*, in case of exceptional measures) the reasoning should be more rigorous[147].

The duty to give reasons requires that the public actor discloses the reasoning followed to reach a decision in a clear and unequivocal way. This means that the reasons must be logical and contain no internal inconsistency that would prevent a proper understanding of the reasons underlying the measure. In terms of extent, the decision-making authority is required to set out the facts and the legal considerations having decisive importance and address all relevant counterarguments. This does not include being required to go into all points of law and fact, also because the statement of reasons is assessed not only

144 CJEU, *Commission of the European Communities v. Chambre syndicale nationale des entreprises de transport de fonds et valeurs (Sytraval) and Brink's France SARL*, C-367/95 P, (1998).
145 CJEU, *Cheil Jedang Corp. v. Commission of the European Communities*, T-220/00, (2003).
146 CJEU, *Elf Aquitaine SA v. European Commission*, C-521/09, (2011).
147 J. Dirutigliano, *Some considerations on the relationship between the right to a reasoned decision and the right to explanation in the proposal of the artificial intelligence act*, The Digital Constitutionalist the Future of Constitutionalism, 2023.

in light of its wording but also its context and all the legal rules governing the matter in question. Beyond these general requirements, the types of reasons to be given depend on the measure in question. Thus, in short, the legal reasoning should consist of three elements: (i) the legal provisions applied; (ii) the relevant facts; and (iii) the decisive considerations in applying the law to the facts, such as specific interpretations of the law or the facts, discretionary choices made, or any other factor that carried weight in the assessment[148].

Considering the context of AI, there are some challenges in implementing this right. A major challenge in relation to the control of AI systems is their opacity. This may be 'intentional opacity', where the AI system's workings remain undisclosed to protect intellectual property rights, 'illiterate opacity', where a system might be understandable, but only to those with specific expert knowledge on coding and computing, or 'intrinsic opacity', where a system is so complex that it is generally not understandable for humans. The opacity of AI systems presents a particular challenge in light of the widely held view that transparency is a necessary condition to enable control over AI systems. Sharing information on the algorithm broadly speaking, the data it relies on, or its internal logic, may be difficult or outright impossible, if an AI system is opaque. Explanation requirements, such as the duty to give reasons, face a similar problem. As a specific form of transparency, they focus on the reasons or justification for an outcome. What is important is not how a decision was reached, but why. Providing an explanation for the decision of an AI system would involve disclosing to a human observer how different factors were weighed by the system to reach a decision and what input was determinative. When a decision is (partially) based on the output of an AI system, the possibility of giving an explanation thus depends on the system itself being able to generate an outcome that

[148] M. Fink – M. Finck, *Reasoned A(I)administration: explanation requirements in EU law and the automation of public administration*, European Law Review, 47, 3, 2022, p. 376 – 392.

is understandable by humans. In other words, some degree of 'explainability' (also 'interpretability') of an AI system is a precondition to produce an explanation. From the perspective of the duty to give reasons, 'illiterate opacity' presents a problem because the workforce of a public body necessarily consists of experts on the specific subject matter at the heart of that bodies' tasks, rather than AI experts. Thus, the public actor will not be able to substantially explain its own decision, if it uses an AI system's recommendation, the justification for which is interpretable only with expert knowledge. Similarly challenging is 'intrinsic opacity', that is the 'black box' problem[149]. As explained also in the previous sections, the image of the black box denotes that the inner workings of complex algorithmic systems, such as deep learning algorithms, can often not be reconstructed by humans. This means that the use of different variables to transform an input into an output cannot be retraced *ex post*. As a result, humans cannot understand how, exactly, these systems reach their conclusions. In part, this stems from the fact that AI relies on correlations, not causation[150]. Indeed, an actual causal relationship between input and output 'may simply not exist, no matter how intuitive such relationships might look on the surface'[151]. This also means they cannot support causal explanations of the kind that underlies the reasons traditionally offered to justify governmental action. In other words, a public authority cannot provide detailed reasons for a decision informed by an algorithm suffering from this type of opacity[152].

Overall, from the perspective of the duty to give reasons, this analysis raises fundamental questions: (i) does a public authority comply with its reason-giving obligations if it simply defers to the recommendation provided by an AI system? (ii) or does EU administrative

[149] Please refer to note 64.
[150] B. Schölkopf, *Causality for Machine Learning*, in (eds.) H. Geffner – R. Dechter – J.Y. Halpern, *Probabilistic and Causal Inference*, New York, 2022, p. 765–804.
[151] C. Coglianese – D. Lehr, *Regulating by Robot*, Georgetown Law Journal, 2017, p. 1147-1157.
[152] M. Fink – M. Finck, *Reasoned A(I)administration: explanation requirements in EU law and the automation of public administration*, European Law Review, 47, 3, 2022, p. 376 – 392.

law require the reasons underlying the AI system's recommendation to be disclosed? (iii) what exactly, in other words, is required from a public authority under the duty to give reasons, when it relies on a recommendation by an AI system?

In this context, it might seem intuitive to assume that in situations where public bodies rely on an AI system's recommendation, their reasoning obligations may be fulfilled by reference to that recommendation. The underlying idea is that what distinguishes this scenario from the one where the authority decides without AI support is added skill, knowledge, or expertise brought in by the system that makes the final decision more reliable and, thus, justifies a lowering of the usual reasoning requirement. Apart from the fact that the underlying assumption that AI systems make better decisions is not necessarily true, two considerations speak against this. First, EU courts have not adopted this particular approach in other areas where EU bodies rely on the findings of expert bodies, such as EU agencies or international institutions. In such scenarios, the decision-making institution nonetheless has to provide an explanation that specifically identifies the concrete reasons why a particular conclusion was reached[153]. This means that even if the use of AI in decision making were considered equivalent to relying on expert advice, providing a justification that simply refers to the fact that an AI system recommended the decision would not meet the requirements under the duty to give reasons. Secondly, and more fundamentally, the perceived correctness of an AI system's output has no bearing on reasoning requirements. This has to do with the specific purpose of the reason-giving obligation. While it is certainly considered to promote accuracy by forcing an authority to lay out and thus ponder the reasons for a decision, its essential function lies in ensuring judicial control over administrative power and enabling private parties to challenge decisions. As explained above, the court may thus well find that a decision is substantively

[153] CJEU, *European Commission v. Kadi*, C-584/10 P, C-593/10 P and C-595/10, (2014); implicitly also CJEU, *Pfizer Animal Health v. Council of the European Union*, T-13/99, (2002).

in accordance with the law but nonetheless annul it for lack of adequate reasons communicated to the addressee. In other words, no matter how good a decision is likely to be from a substantive perspective, concrete reasons have to be provided so that the decision can be challenged and reviewed. Thus, it is safe to say that 'the machine said so' will not be sufficient under the law. This also means that, at least where the AI recommendation is the only or main factor affecting the direction of a decision, the public authority that takes the decision needs to understand the reasons that determined the AI system's output itself, in order to comply with its reason-giving obligations[154].

Therefore, a degree of explainability of the AI system used is necessary to produce an explanation that complies with the requirements under EU administrative law. Beyond this, the fact that AI is used may actually be a reason to increase the decision-maker's reasoning obligations. As explained above, in areas of high factual (economic, technical, or scientific) complexity, where the public administration enjoys a wider margin of appraisal, the EU courts have required them to more thoroughly reason their decisions *vis-à-vis* those affected. This is to counterbalance the fact that the Union judge may not, sometimes even cannot, fully review the substantive legality of decisions of the (expert) administration. Two considerations are relevant in this respect. First, AI offers the possibility to process vast amounts of information, a tool needed in areas of factual complexity. Thus, AI systems may often be used in areas already characterized by increased reasoning obligations. Secondly, even when not actually used in an area of factual complexity, some AI systems, in particular those that suffer from the opacity described earlier, pose similar constraints to the court's ability to conduct full substantive review. To balance this limit to full substantive review, the court might consider increased scrutiny of reasoning obligations, similar to the approach adopted in areas of wide

[154] M. Fink – M. Finck, *Reasoned A(I)administration: explanation requirements in EU law and the automation of public administration*, European Law Review, 47, 3, 2022, p. 376 – 392.

discretion and high complexity. In addition, it can be argued that the court should apply a higher explanation requirement when decision-making is automated to counter-act the risk of 'automation bias'. Automation bias refers to the phenomenon that humans ascribe a certain authority to outcomes suggested by automated processes that lead them to neglect other available information or counter-indications[155]. Thus, even where an authority can give a justification, it may in substance have 'blindly' trusted the outcome suggested by the AI system[156]. Preventing automation bias might necessitate additional safeguards, including, for instance, requiring the public authority that relies on AI to communicate how other available information or alternative outcomes were considered in reaching a decision[157].

In sum, the duty to give reasons in EU administrative law provides a useful starting point to think about the rights individuals have to obtain an explanation in cases where public authorities rely on AI to reach their decisions. It makes blind 'rubberstamping' of AI recommendations more difficult by requiring explanations that go beyond a simple 'the machine said so'. A case might even be made for more extensive reasoning obligations for public authorities when they use AI in their decision-making to make up for the possible limits this poses to substantive legality review and to counter-act the risk of automation bias[158]. In any case, at least where AI is the only or main factor affecting the direction of a decision, a degree of explainability would seem necessary to produce an explanation that complies with the requirements under EU administrative law. In this context, secondary

[155] Common examples include car accidents that are caused by a human driver relying on a route suggested by a driving assistant even when real-life factors speak against it.
[156] M. Fink – M. Finck, *Reasoned A(I)administration: explanation requirements in EU law and the automation of public administration*, European Law Review, 47, 3, 2022, p. 376 – 392.
[157] Other possible safeguards may include introducing a 'four eyes principle' where a second official has to approve a decision and the reasoning behind it before it enters into force.
[158] There are examples of legal systems that do contain higher explanation threshold when the decision-making process is automated. Noting this for the French legal system see A. Bibal – M. Lognoul – A. De Streel – B. Frénay, *Legal Requirements on Explainability in Machine Learning*, Artificial Intelligence and Law, 2021, p. 154–155.

law could play an important role in specifying the exact extent to which the public authority's reasoning may rely on an AI system's recommendation and what that means for the degree of explainability required from the system[159].

2.4.4 The accountability of the public actor

Art. 41(3) states that:

> 'every person has the right to have the Union make good any damage caused by its institutions or by its servants in the performance of their duties, in accordance with the general principles common to the laws of the Member States'.

The aim of this provision is the individual's protection against illegal and harmful conduct of the public actors. As such, the public institution is responsible for lawless actions and neglect, if an obligation to carry out specific actions has been in place[160].

The action is illegal when it violates superior law (*e.g.*, fundamental rights, duty of care and the principle of good administration), which means that it infringes provision intended to protect individuals' rights. Moreover, it must be sufficiently flagrant violation of a superior law. Particularly, an infringement of the law is sufficiently serious if public institutions manifestly and gravely disregard the limits on their discretion[161]. Moreover, the public actor is liable for damages caused by its institutions or its officials in the performance of their duties. Performance of the duties constitutes every action of the institutions and servants 'which, by virtue of an internal and direct relationship, are the necessary extension of the tasks entrusted to the institutions'[162]. Illegal conduct of institutions and servants must

[159] M. Fink – M. Finck, *Reasoned A(I)administration: explanation requirements in EU law and the automation of public administration*, European Law Review, 47, 3, 2022, p. 376 – 392.
[160] J. Maliszewska-Nienartowicz, *Porządek prawny Unii Europejskiej* (The Legal Order of the European Union), Toruń, 2005, 238.
[161] CJEU, *Casper Koelman v. Commission of the EC*, T-575/93, (1996) states that 'Where the institution in question has only a considerably reduced or even no discretion, the mere infringement of […] law may be sufficient to establish the existence of a sufficiently serious breach'.
[162] CJEU, *Claude Sayag and S.A. Zurich v. Jean-Pierre Leduc, Denise Thonnon and S.A. La Concorde*, C-9/69, (1969).

cause damage to individuals. Applicants must provide the court with evidence to establish the fact and the extent of the loss which he or she claims to have suffered[163]. The damage must be done *de facto* or must be foreseeable with sufficient certainty.

In this context, it appears that the public actor is required to be responsible for the decisions made using algorithms, as accountability refers to the obligation to explain and justify the decisions adopted on the basis of AI. Through accountability the public actor is expected to justify its use of algorithms in decision-making and explain how they work, what data they use, what results are expected and what results have been obtained. Accountability facilitates the monitoring of activities performed by public administrations using algorithms. It also allows errors and biases to be detected. Ultimately, accountability strengthens public trust in public administrations in relation to their use of algorithms for decision-making.

In this context, in order to hold the public actor accountable there should be total transparency on AI uses, their purposes and operation, as well as the data used so that public employees are aware of all these factors, and so that they can, ultimately, be monitored by citizens. In addition to this, it should be considered that the public actor can also be accountable for decisions taken by algorithms using empirical evidence that allows the government to measure the performance of algorithms and demonstrate the benefits and impacts with which to justify their decisions, without the need for an explanation of the algorithm[164].

In this normative framework, it is necessary to investigate whether secondary legislation provides for specific mechanisms that are able to make the public actor accountable for the decision taken through AI technologies.

---

[163] CJEU, *Casper Koelman v. Commission of the EC*, T-575/93, (1996).

[164] A. C. Martínez, *How can we open the black box of public administration? Transparency and accountability in the use of algorithms,* Revista catalana de dret public, 2019, p. 13-28.

2.5 The right to an effective remedy

One of the Charter most relevant provisions is Art. 47, which provides for the right to an effective remedy and to a fair trial. It states that:

> ‘Everyone whose rights and freedoms guaranteed by the law of the Union are violated has the right to an effective remedy before a tribunal in compliance with the conditions laid down in this Article.
>
> Everyone is entitled to a fair and public hearing within a reasonable time by an independent and impartial tribunal previously established by law.
>
> Everyone shall have the possibility of being advised, defended and represented.
>
> Legal aid shall be made available to those who lack sufficient resources in so far as such aid is necessary to ensure effective access to justice’.

The right to an effective judicial remedy is deeply engrained in the constitutional order of the EU. The CJEU describes the origins and nature of this right as being a provision which is a general principle of EU law. It is linked to obligations in the first subparagraph of Art. 19(1) TEU under which European courts ‘shall ensure that in the interpretation and application of the Treaties the law is observed’. Under the second paragraph of the same article, Member States shall, through their national courts, provide ‘remedies sufficient to ensure effective legal protection in the fields covered by Union law’. The CJEU has described this in one general formula, stating as follows:

> ‘The principle of the effective judicial protection of individuals’ rights under EU law, referred to in the second subparagraph of Art. 19(1) TEU, is a general principle of EU law stemming from the constitutional traditions common to the Member States, which has been enshrined in Art. 6 and 13 of the European Convention for the Protection of

Human Rights and Fundamental Freedoms […] and which is now reaffirmed by Art. 47 of the Charter'.

Particularly, as consistently stated by the CJEU, the right to an effective remedy and to a fair trial 'comprises various elements; in particular, the rights of the defense, the principle of equality of arms, the right of access to a tribunal and the right to be advised, defended and represented'[165]. Tribunal is not defined in this provision, but the relevant criteria are those which the CJEU takes into account when determining whether a body is a court or tribunal for the purposes of Art. 267 TFEU[166]. Another core element of Art. 47(1) is that remedies for violations of rights and freedoms guaranteed by EU law need to be effective and real. This implies certain rights of defense.

The algorithmic age demands a more refined remedial framework beyond judicial remedies. Indeed, the emerging EU digital acquis establish an extensive array of *ex ante* accountability mechanisms, including impact assessments, continuous reporting and informing duties, or horizontally applicable common technical standards. Not all of these mechanisms constitute direct remedies. The latter can take different forms in the chain of remedial actions, culminating with the individual's right to a remedy before the court. Combinations of administrative and judicial review mechanisms are widely spread across EU policy areas. As the CJEU clarified in the case of Puškár[167], rules prescribing an obligation to first exhaust administrative mechanisms before seeking a judicial review constitute legitimate limits on the right to an effective judicial protection. These rules

[165] CJEU, *Otis*, C-199/11, (2012); CJEU, *Sacko* C-348/16, (2017); CJEU, *EG*, C-662/17, (2018).
[166] CJUE, *D e Coster*, C -17/00, (2001); CJEU, *Torresi*, C -58/13, (2014); CJEU, *Associação Sindical dos Juízes Portugueses*, C-64/16, (2018); CJEU, *Commission v. Poland*, C-619/18, (2019). On the concept of a court or tribunal and the autonomy of EU law, CJUE, *Achmea*, C-284/16, (2018).
[167] CJUE, Case C-73/16, *Puškár*, (2017).

reduce the burden already placed on the courts, thus ultimately reinforce efficiency rather than weakening the remedies[168].

With regard to the application of the right to an effective remedy within the use of AI by the public actor, it has to be highlighted that the aforementioned right covers also decisions taken with the support of AI technologies. Particularly, procedural fairness means that individuals should have the ability to contest and seek effective redress against decisions made by AI systems and by the humans operating them[169]. In this context, the consequences of AI applications on this specific right may be specifically alarming[170]. Particularly, remedy comprises effective reparation, appropriate accountability for those responsible, as well as measures to prevent recurrences. This means that, at all stages of design and deployment of AI, it must be clear who bears responsibility for its operation. Particularly, clarity is required where the division of responsibilities lies between the developer of an AI system, the deployer of the system (*i.e.*, the public actor), and the user. Consequently, users of AI systems will need adequate understanding or assurance as to how those systems work. Complainants need to know how to complain and to whom, and to be confident that their complaint will be addressed in a timely manner. Remedy relies on transparency and explainability as complainants should have enough information to understand how a decision about them was made, and the role and operation of AI in the decision-making process. They may need access to data on how the AI was designed and tested, how it was intended to operate and how it operated in specific cases, as well as information on the role of human decision-making or oversight in the process[171].

---

[168] G. De Gregorio – S. Demková, *The Constitutional Right to an Effective Remedy in the Digital Age: A Perspective from Europe*, European Yearbook of Constitutional Law 2023: Constitutional Law in the Digital Era, 2024, p. 223-254.

[169] G. Sartor, *Artificial Intelligence and Human Rights: Between Law and Ethics*, Maastricht Journal of European and Comparative Law, 27, 2020, p. 705–719.

[170] S. De Heer, *Artificial Intelligence and the Right to an Effective Remedy*, in (eds.) A. Quintavalla – J. Temperman, *Artificial Intelligence and Human Rights*, Oxford, 2023, p. 294.

[171] K. Jones, *AI governance and human rights, Resetting the relationship*, Chatman House, 2023.

However, the question arises how an individual ought to seek redress for impaired substantive rights with limited or no knowledge on how these generic AI applications operate[172]. Since the use of AI by the public actor challenges the right to an effective remedy in different ways[173], it is necessary to observe the remedies adopted by the secondary legislation[174].

3. Regulation of AI: from the starting point to the landing point

As pointed out in the previous sections, the use of AI by the public actor, if not proper regulated, can result in an excessive restriction of the exercise of certain citizens' fundamental rights. For this reason, it is necessary to analyze secondary legislation in order to understand whether and how the European legislator has sought to solve the problems related to the guarantee of the right to privacy and personal data protection, the right to non-discrimination, the right to have a good administration, as well as the right to have an effective remedy.

In this context, the GDPR marked the beginning of a revolution that has seen Europe as the first to impose a set of requirements needed for processing personal data (and therefore, for the use of the systems that consequently process personal data). As such, considering the large amount of personal data processed through AI also in the context of

[172] S. De Heer, *Artificial Intelligence and the Right to an Effective Remedy*, in (eds.) A. Quintavalla – J. Temperman, *Artificial Intelligence and Human Rights*, Oxford, 2023, p. 294.

[173] European Union Agency for fundamental rights, *Getting the Future Right Artificial Intelligence and Fundamental Rights*, Luxembourg: Publications Office of the European Union, (December 2020).

[174] As examined in the following sections, Art. 85 of the AI Act states that 'Without prejudice to other administrative or judicial remedies, any natural or legal person having grounds to consider that there has been an infringement of the provisions of this Regulation may submit complaints to the relevant market surveillance authority'. While this is an improvement over the original proposal, it falls short of guaranteeing procedural rights comparable to those arising from Art. 47 EUCFR. Indeed, it is unclear at what stage such a complaint can be realistically expected to be filed with the relevant authority, but it seems unlikely that, even with increased transparency (*e.g.*, of the fundamental rights impact assessment), individuals will have access to the information required to assess non-compliance other than in relation to formal aspects of the AI Act's obligations. Material deviations from the relevant guarantees and, in particular, violations of fundamental rights not captured or incorrectly evaluated in the relevant impact assessment will most likely only be discovered after the fact. This does not do much to address the issue of the (too late) timing and (in)effective access to remedies in the context of mass administrative decision-making.

exercising public functions, GDPR's provisions are extremely relevant. This also taking into account that, with the AI Act, the GDPR constitutes one of the most comprehensive regulations about new technologies. Therefore, its analysis is presented in section 3.1.

Moreover, since the application of GDPR to AI in these last years has left open many interpretative questions, case law is also described in section 3.2. Indeed, due to the absence of an *ad hoc* legislation, the jurisprudential body – as it often happens – has had to resolve several issues connected to the use of AI and the exercise of competing fundamental rights.

Finally, section 3.3 examines the very recent text of the IA Act, as it can be seen as the 'landing point' of the AI regulation after a lengthy approval process by Europe. This analysis aims to point out what regulatory and technical tools the European legislator has decided to adopt, as well as to examine whether the AI Act gains the aim of 'facilitating the protection of natural persons, undertakings, democracy, the rule of law and environmental protection, while boosting innovation and employment and making the Union a leader in the uptake of trustworthy AI'[175].

3.1 The GDPR: The starting point

In the last few years, a huge debate has been related to the possibility of applying the GDPR to the use of AI, as in 2016 this Regulation was not conceived for this new type of technology. Indeed, considering that many of the AI Act's provisions will be applicable from 2026, in the meanwhile it must be ensured that the development and deployment of AI tools takes place in a socio-technical framework – inclusive of technologies, human skills, organizational structures, and norms – where individual interests and the social good are preserved and enhanced. To provide regulatory support for the creation of such a

[175] Recital 2 of the AI Act.

framework, the sectoral regulations involved include first of the data protection laws and specifically the GDPR. As such, this Regulation can be seen as the starting point of the analysis of the relevant laws within the European Union legal order.

In this context, AI systems are not explicitly mentioned in the GPDR, even if many provisions in the Regulation are relevant to them, and some are indeed challenged by the new ways of processing personal data that are enabled by these systems. Indeed, there is a tension between the traditional data protection principles – purpose limitation, data minimization, the special processing of 'sensitive data', the limitation on automated decisions – and the full deployment of the power of AI and big data. The latter entails the collection of vast quantities of data concerning individuals and their social relations and processing such data for purposes that are not fully determined at the time of collection. However, it is necessary to assess whether there are ways to interpret, apply, and develop the data protection principles consistently with the beneficial uses of AI and big data[176]. Therefore, the following sections examine the most important provisions of the GDPR[177] related to the use of AI.

3.1.1 Applicability of the definitions of the GDPR to AI

According to the definition given in Art. 4(1) of the GDPR the term 'personal data' includes all information concerning an identifiable natural person, or the 'data subject'. 'Identifiable

[176] Panel for the Future of Science and Technology EPRS, *The impact of the General Data Protection Regulation (GDPR) on Artificial intelligence*, European Parliamentary Research Service Scientific Foresight Unit (STOA), (June 2020).

[177] Before presenting the analysis of the GDPR, it is important to make an important premise. Indeed, as far as the EU administration is concerned, the GDPR does not actually apply, as the relevant law is the Regulation (UE) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC (EUDPR), which regulates how EU institutions, bodies and agencies should process personal data. However, the provisions related to the automated decision-making processing provided for in the GDPR are identically replicated in the EUDPR. Indeed, Art. 22 and Recital 71 of the GDPR are equivalent to Art. 24 and Recital 43 of the EUDPR. This means that GDPR and EUDPR share a common formulation. This feature allows to present the analysis just in relation to the GDPR's provisions.

person' is a person that can be identified by reference to certain characteristics that comprise that person's identity, such as name, location data, psychological and genetic traits, economic qualities, etc. As for the latter, Art. 4(2) of the GDPR defines 'processing' as operations that are performed on personal data, including the recording, storing, alteration, use, destruction, etc. by automated or other means. Consequently, if a public administration, when making decisions concerning public services, employs an automated system that analyzes a person's information, and compares that data to similar data from other citizens in the same situation, such conduct indeed falls within the scope of the GDPR. Indeed, section 4 of the same article devotes specific attention to 'profiling', a term which applies more precisely to the many applications of AI technology and is defined as any form of automated processing conducted for the purpose of evaluating certain aspects or a particular state pertaining to a natural person, such as work performance, economic situation, interests, behavior, etc.

However, at the same time, Robert van den Hoven van Genderen does not find this (or any) provision of the GDPR to govern AI specifically, as it argues that '[s]urprisingly, there is no vision on the use of AI and robotics in the GDPR. These concepts are nowhere to be found in the text or recitals'[178]. This disagreement, and the general nature of the provision related to profiling, can be explained in technical terms. For an automated decision to be made, for example, regarding the allocation of public jobs following a ranking list, the use of AI techniques is not required as the automated decision processing used has 'simpler' characteristics and it does not take discretionary decisions. Using AI, on the other hand, provides for a much more accurate result due to the ability of modern computers to process vast amounts of data of all kinds in real time and to establish patterns between those pieces of data. Simple algorithms would satisfy the characteristics

[178] R. van den Hoven van Genderen, *Privacy and Data Protection in the Age of Pervasive Technologies in AI and Robotics*, European Data Protection Law Review, 3, 2017, p. 338 – 352.

of automated decision-making under Art. 4(4) and would make the satisfaction of the requirement of explainability an easy task. However, an AI system has much more complicated characteristics, and this complexity does not seem to be specifically addressed in the GDPR.

Considering this premise and due to the lack of specific provisions, it is clear the need to explore the legal issues surrounding the deployment of AI systems under the regime established by the GDPR[179].

3.1.2 Art. 22: The regulation of the automated individual decision-making, including profiling

To the above declared end, it is necessary to start from Art. 22 of the GDPR which states that:

> ‘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.
>
> Paragraph 1 shall not apply if the decision:
>
> a) is necessary for entering into, or performance of, a contract between the data subject and a data controller;
> b) is authorized by Union or Member State law to which the controller is subject and which also lays down suitable measures to safeguard the data subject’s rights and freedoms and legitimate interests; or
> c) is based on the data subject’s explicit consent.
>
> In the cases referred to in points (a) and (c) of paragraph 2, the data controller shall implement suitable measures to safeguard the data subject’s rights and freedoms and

[179] A. Kesa – T. Kerikmäe, *Artificial Intelligence and the GDPR: Inevitable Nemeses?*, TalTech Journal of European Studies, 10(3), 2020, p. 68-90.

legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision.

Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Article 9(1), unless point (a) or (g) of Article 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place'.

In accordance with Art. 22 of the GDPR, data subjects have the right not to be subjected to a decision based solely on automated processing which produces legal effects concerning him or her or, similarly, significantly affects him or her.

There are three exceptions to this rule: (i) explicit consent; (ii) when such processing is necessary for entering into or performing a contract, where the controller and data subject are parties, or; (iii) it is authorized by European Union or Member State law. In every case suitable safeguard must be implemented and, at least in the context of consent or contract, the data subject must be able to obtain human intervention. Further restrictions are in place for special categories of personal data, as decision based solely on automated processing is only admissible when there is specific consent or reasons of substantial public interest (and, of course, pursuant to abovementioned suitable safeguards). Recital 71 complements this provision stating that:

'such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision'[180].

[180] T. S. Cabral, *AI and the Right to Explanation: Three Legal Bases under the GDPR*, in (eds.) D. Hallinan – R. Leenes – P. De Hert, *Data Protection and Privacy: Data Protection and Artificial Intelligence*, Oxford, 2021, p. 29–56.

Considering both Art. 22 and Recital 71, their analysis in the context of the use of AI are presented below (section 3.1.2.1 – section 3.1.2.4).

3.1.2.1 The prohibition of automated decision-making processing

The first paragraph of Art. 22 provides for a general right not to be subject to completely automated decisions significantly affecting the data subject.

Even though this provision refers to a right, it rather introduces a prohibition upon data controllers as automated decisions affecting data subjects are prohibited, unless they fit in with one of the exceptions provided in paragraph 2[181].

On this merit, according to Article 29 Working Party (WP29), as a rule, there is a general prohibition on fully automated individual decision-making, including profiling that has a legal or similarly significant effect[182]. For the application of the prohibition established by Art. 22(1), four conditions are needed: (i) a decision must be taken; (ii) it must be solely based on automated processing; (iii) it must include profiling; (iv) it must have legal or anyway significant effect.

As regards the first condition relating to the existence of a decision, it should be noted that the concept of 'decision', within the meaning of Article 22(1) of the GDPR, is not defined by the Regulation. However, it is apparent from the very wording of that provision that this concept refers not only to acts which produce legal effects concerning the person at issue but also to acts which similarly significantly affect him or her. The broad scope of the concept of 'decision' is confirmed by Recital 71 of the GDPR, according to which a decision evaluating personal aspects relating to a person, to which that person should

[181] I. Mendoza – L. A. Bygrave, *The Right Not to Be Subject to Automated Decisions Based on Profiling,* in (eds.) T. Synodinou – P. Jougleux – C. Markou – T. Prastitou, *EU Internet Law: Regulation and Enforcement*, 2017.

[182] WP29, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, (October 3rd, 2017).

have the right not to be subject, ‘may include a measure’ which either produces ‘legal effects concerning him or her’, or, ‘similarly significantly affects him or her’. Under that recital, the term ‘decision’ covers, for example, the automatic refusal of an online credit application or e-recruiting practices without human intervention[183].

The second condition requires that humans do not exercise any real influence on the outcome of a decision-making process, even though the final decision is formally ascribed to a person[184]. On this point, there is some controversy about what should be interpreted as ‘based solely on automated decision’, that should be addressed. In trilogue[185] negotiations the European Parliament had proposed a broader formulation which included ‘solely or predominantly’. The apparently restrictive final choice has been the source of some degree of confusion. Did the European legislator intend to enshrine a regime where the mere fact that a human participating in the decision-making excludes the application of Art. 22? The use of the adverb ‘solely’ implies that a decision does not involve anyone or anything else. That is to say it is exclusively derived from the machine’s decision. ‘Solely’ implies, realistically, that the power to make the decision rests exclusively in the program. If the person who is supervising it does not have the power to override or does not understand its functioning in a manner that would allow him/her to

[183] CJEU, *Schufa*, C-634/21, (2023), para. 44-45. The CJEU ruled that a credit reference agency engages in automated individual decision-making when it creates credit repayment probability scores as a result of automated processing and where lenders rely heavily on these scores (to establish, implement, or terminate contracts). This means the obligation to comply with Art. 22 of the GDPR falls on the credit reference agency rather than just on the lender. Art. 22 only allows the use of automated individual decision-making in limited situations, *e.g.*, contractual necessity, specific justifications under EU or member state law, or explicit consent. Where individual decision-making is allowed, protections must be in place – including ensuring that individuals can obtain human intervention, express their views, and challenge decisions made about them. Regulators and other stakeholders have stated the decision will have ramifications for all automated decision-making services and providers offering predictive AI tools. Therefore, under this case, the notion of ‘decision’ is broad as it includes also acts that may affect the individual in various ways, including the calculation of a credit score.

[184] Panel for the Future of Science and Technology EPRS, *The impact of the General Data Protection Regulation (GDPR) on Artificial intelligence*, European Parliamentary Research Service Scientific Foresight Unit (STOA), (June 2020).

[185] T. S. Cabral, *A Short Guide to the Legislative Procedure in the European Union’*, UNIO – EU Law Journal, 2020, p. 161–80.

detect errors[186] and correct them, said power still sits solely with the program and, thereby, Art. 22 is applicable[187]. In its Guidelines on automated individual decision-making and profiling[188] (Guidelines), the WP29 clarified its opinion about what it understood by human involvement. Particularly, it considers that data controllers cannot fabricate human involvement to avoid the application of Art. 22. According to the Guidelines 'to qualify as human involvement, the controller must ensure that any oversight of the decision is meaningful, rather than just a token gesture. It should be carried out by someone who has the authority and competence to change the decision. As part of the analysis, they should consider all the relevant data'[189].

The third condition requires that the automated processing determining the decision includes profiling. A different interpretation could be suggested by the comma that separates 'processing' and 'including profiling' in Art. 22(1), which seems to indicate that profiling only is an optional component of the kind of automated decisions that are in principle prohibited by Art. 22(1). However, the first interpretation (the necessity of profiling) is confirmed by Recital 71, according to which the processing at stake in the Regulation of automated decision must include profiling. Such processing includes 'profiling' that consists of any form of automated processing of personal data evaluating the personal aspects relating to a natural person, in particular to analyze or predict aspects concerning the data subject's performance at work, economic situation, health, personal preferences or interests, reliability or behavior, location or movements.

[186] Indeed, the person should be able to detect errors, at least, at the level of someone who is reasonably familiarized with the program and is reasonably diligent.

[187] The analysis should be, as with the establishment of the legal position of the parties as data processor or controller, based on the situation *de facto* and not *de jure*.

[188] WP29, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, (October 3rd, 2017).

[189] T. S. Cabral, *AI and the Right to Explanation: Three Legal Bases under the GDPR*, in (eds.) D. Hallinan - R. Leenes – P. De Hert, *Data Protection and Privacy: Data Protection and Artificial Intelligence*, Oxford, 2021, p. 29–56.

The fourth condition requires that the decision produces legal effects concerning the data subject or similarly significantly affects him or her[190]. The GDPR offers no definition about what should be considered as a decision producing legal or similarly significant effects. Recital 71 presents the example of 'automatic refusal of an online credit application or e-recruiting practices without any human intervention' but it is not possible to find further clarification of the legal text. Again, the input of WP29 is invaluable to provide some degree of clarity. In accordance with its Guidelines, a decision shall be considered having legal effects when it affects a person's legal rights, legal status, or rights under a contract. However, and as reasonable as this seems at first glance, one has to note that it can be rather troublesome. What is an adverse effect on the legal rights of the data subject? It seems that the definition of an effect that is similar is slightly more difficult and must be accessed on a case-by-case basis. According to the Guidelines, the threshold should be that those effects should affect the data subject in a similar manner as the legal effects. WP29 establishes three criteria: the decision must have the potential to: (i) 'significantly affect the circumstances, behavior or choices of the individuals concerned'; (ii) 'have a prolonged or permanent impact on the data subject' or; (iii) 'at its most extreme, lead to the exclusion or discrimination of individuals'. A similarly significant effect can arise from situations created by third parties and not necessarily by the data subject[191].

Following the elements described, it is possible to assert that many decisions made by AI systems could fall under the scope of Art. 21(1). Indeed, the use of AI makes it more likely that a decision will be based 'solely' on automated processing. This is due to that humans may not have access to all the information that is used by AI systems and may not have

[190] WP29, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, (October 3rd, 2017).
[191] T. S. Cabral, *AI and the Right to Explanation: Three Legal Bases under the GDPR*, in (eds.) D. Hallinan – R. Leenes – P. De Hert, *Data Protection and Privacy: Data Protection and Artificial Intelligence*, Oxford, 2021, p. 29–56.

the ability to analyze and review the way in which this information is used. It may be impossible, or it may take an excessive effort to carry out an effective review – unless the system has been effectively engineered for transparency, which in some cases may be beyond the state of the art. Thus, especially when a large-scale opaque system is deployed, humans are likely to merely execute the automated suggestions by AI, even when they are formally in charge. Moreover, human intervention may be prevented by the costs-and-incentives structure in place[192].

That being ascertained, it is necessary to analyze the other elements of Art. 22 as, particularly, the exceptions to the prohibition of automated decision-making processing and the related safeguards.

3.1.2.2 Exceptions to the prohibition of automated decision-making processing and safeguards

Paragraph 2 of Art. 22 provides for three broad exceptions to paragraph 1. It states that the prohibition on automated decision-making does not apply when the processing upon which the decision is based a) is necessary for entering into, or performance of, a contract between the data subject and the data controller; b) is authorized by Union or Member State law to which the controller is subject, and which also lays down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests; or c) is based on the data subject's explicit consent.

Based on the broad exception of item a), automated decision-making can be enabled in key areas such as the exercise of public functions. However, for the exception to be applicable, decisions based solely on automated processing must be 'necessary'. Such

[192] Panel for the Future of Science and Technology EPRS, *The impact of the General Data Protection Regulation (GDPR) on Artificial intelligence*, European Parliamentary Research Service Scientific Foresight Unit (STOA), (June 2020).

necessity, for instance, may depend on the high number of cases to be examined. The necessity of using AI in decision-making may also be connected to AI capacities to outperform human judgement. In this connection, it can be asked whether human involvement can still contribute to a stronger protection of data subjects, or whether the better performance of machines makes human intervention redundant or dysfunctional[193]. Furthermore, it could be argued that such an exception leaves ample room for parties' power of negotiating, which, however, in most cases is not equally spread between them. Specifically in the public field, it is clear that that the public actor could have a stronger position in negotiations, being able to force the other party to the use of automated decision-making.

Moreover, automated decision-making, including profiling, could potentially take place under Art. 22(2)(b) as Union or Member State law authorizes it. The relevant law must also lay down suitable measures to safeguard the data subject's rights and freedoms and legitimate interests. Recital 71 says that this could include the use of automated decision-making defined in Art. 22(1) for monitoring and preventing fraud and tax-evasion, or to ensure the security and reliability of a service provided by the controller. In the public field, this exception is particularly relevant because it allows the public actor to create *ad hoc* regulations for the use of automated decision-making including profiling, which thus avoids the strict scrutiny of Art. 22 GDPR.

Lastly, outside of the domain of contract and legal authorization, consent may provide a basis for automated decision-making according to Art. 22(2)(c). However, the conditions for valid consent do not always seem appropriate. Consider, for instance, the case in which the public actor uses an automated method for classifying (profiling) applicants to determine their need and consequently allocate certain benefits to them. In such a case,

[193] *Ibid*.

it is very doubtful that an applicant's consent may be viewed as free (as not consenting would entail being excluded from the benefit)[194].

In any case, when Art. 22(2)(a) and (c) applies – *i.e.*, when the automated decision is necessary to contract or explicitly consented – the controller is subject to additional information obligations under Art. 13(2)(f) and 14(2)(g) of the GDPR. Secondly, the data subject enjoys, under Art. 15(1)(h) of the GDPR, the right to obtain from the controller, in particular, 'meaningful information about the logic involved, as well as the significance and the envisaged consequences of such processing for the data subject'. Moreover, Art. 22(3) requires suitable safeguard measures to be applied. Particularly, the data controller shall implement suitable measures to safeguard the data subject's rights and freedoms and legitimate interests, at least the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision. According to WP29, some of these measures concern risk reduction, which are, for instance, quality assurance checks, algorithmic auditing, data minimization, anonymization or pseudonymization, and certification mechanisms[195]. Such measures should ensure that the requirements set forth in Recital 71 – concerning acceptability, accuracy and reliability – are respected as the controller should: (i) use appropriate mathematical or statistical procedures for the profiling; (ii) implement technical and organizational measures appropriate to ensure, in particular, that factors which result in inaccuracies in personal data are corrected and the risk of errors is minimized; (iii) secure personal data in a manner that the potential risks involved for the interests and rights of the data subject are taken into account and that prevents, *inter alia*, discriminatory effects on natural persons on the basis of racial or ethnic origin, political opinion, religion or beliefs, trade union

[194] *Ibid*.
[195] WP29, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, (October 3rd, 2017), p. 32.

membership, genetic or health status or sexual orientation, or that result in measures having such an effect[196]. According to WP29, the input data must be shown to not be 'inaccurate or irrelevant, or taken out of context', and to not violate 'the reasonable expectations of the data subjects', in relation to the purpose for which the data was collected[197]. In approaches based on machine learning, this should apply not only to the data concerning the person involved in a particular decision, but also to the data in a training set, where the biases built into the training set may affect the learned algorithmic model, and hence the accuracy of the system's inferences. Other measures pertain to the interaction with the data subjects, such the right to obtain human intervention and the right to challenge a decision[198].

Therefore, the applicability of all these safeguards is necessary to ensure that the processing of personal data within AI systems is made in full compliance with the requirements of the GDPR, ultimately aimed at protecting individuals' interests and rights. However, all the other information potentially used by the AI systems (*i.e.*, not strictly personal data) is out of the scope of this discipline, thus creating a normative gap in relation to the regulation of AI systems.

---

[196] See also CJEU, *Schufa*, C-634/21, (2023): 'In the light of recital 71 of the GDPR, such measures must include, in particular, the obligation for the controller to use appropriate mathematical or statistical procedures, implement technical and organizational measures appropriate to ensure that the risk of errors is minimized and inaccuracies are corrected, and secure personal data in a manner that takes account of the potential risks involved for the interests and rights of the data subject and prevent, inter alia, discriminatory effects on that person. Those measures include, moreover, at least the right for the data subject to obtain human intervention on the part of the controller, to express his or her point of view and to challenge the decision taken in his or her regard. It is also important to note that, in accordance with the settled case-law of the Court, any processing of personal data must, first, comply with the principles relating to the processing of data established in Article 5 of the GDPR and, secondly, in the light, in particular, of the principle of the lawfulness of processing, laid down in Article 5(1)(a), satisfy one of the conditions of the lawfulness of the processing listed in Article 6 of that regulation (judgment of 20 October 2022, Digi, C-77/21, EU:C:2022:805, paragraph 49 and the case-law cited). The controller must be able to demonstrate compliance with those principles, in accordance with the principle of accountability set out in Article 5(2) of that regulation (see, to that effect, judgment of 20 October 2022, Digi, C-77/21, EU:C:2022:805, paragraph 24)'.

[197] WP29, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, (October 3rd, 2017), p. 32.

[198] Panel for the Future of Science and Technology EPRS, *The impact of the General Data Protection Regulation (GDPR) on Artificial intelligence*, European Parliamentary Research Service Scientific Foresight Unit (STOA), (June 2020).

3.1.2.3 Automated decision-making and sensitive data

Lastly, it should be noted that Art. 22(4) introduces a prohibition, limited by an exception, to ground automated decisions on 'sensitive data', *i.e.*, the special categories set out in Art. 9(1). Decisions referred to in paragraph 2 shall not be based on special categories of personal data referred to in Art. 9(1), unless letter a) or g) of Art. 9(2) applies and suitable measures to safeguard the data subject's rights and freedoms and legitimate interests are in place. The exception concerns the cases in which the data subject has given explicit consent (Art. 9(2)(a)) or processing is necessary for reason of public interest (Art. 9(2)(g)). The role of the data subject's consent needs to be clarified since consent does not exclude that the method used for the decision is inacceptable (*e.g.*, when it is discriminatory).

However, AI challenges the prohibition of processing sensitive data. First of all, sensitive data can be (probabilistically) inferred from non-sensitive data. For instance, sex orientation can be inferred from a data subject's Internet activity, likes or even facial features. In this case the inference of sensitive data should count as a processing of sensitive data and therefore would have to be considered unlawful unless the conditions under Art. 9 are met. Secondly, non-sensitive data can work as proxies for sensitive data correlated to them, even though the latter are not inferred by the system. For instance, the place of residence can act as a proxy for ethnicity. In this case, unlawful discrimination may take place[199].

Therefore, the automated decisions processing of sensitive data, especially if made through AI technologies, has to comply with specific rules as set out in Art. 22(4).

[199] *Ibid*.

3.1.2.4 The right to an ex-post explanation

Once one of the exceptions provided for in Art. 22(2) applies, the GDPR introduces a right to an explanation, which is currently envisaged in Recital 71 as it establishes that:

> 'In any case, such processing should be subject to suitable safeguards, which should include specific information to the data subject and the right to obtain human intervention, to express his or her point of view, to obtain an explanation of the decision reached after such assessment and to challenge the decision'.

The aim of this right is to make decisions understandable so that they can be challenged and, thus, be subject to some form of control. This would guarantee important fundamental rights (*e.g.*, the right to a good administration, the right to an affective remedy). Moreover, this would also correspond to one of the core objectives of data protection law and particularly to address power asymmetries between data subjects and data controllers. Considering its potential importance, this Recital has generated a huge debate as (i) the right to explanation is added in a non-legally binding provision; (ii) it is not clear whether or not a right to an *ex-post* explanation connected to the automated decision processing actually exists.

On the one hand, some consider that data protection law requires 'human-intelligible explanations of algorithmic decision-making'. Therefore, it should be possible to interpret the combination of Art. 22 and Recital 71 as the proof that a right to *an ex-post* explanation is established. Others understand the right to an explanation as just an *ex-ante* right to information about the use of AI. Indeed, only an *ex-ante* explanation of system functionality is explicitly required by Art. 13(2)(f) and 14(2)(g), which impose notification duties that precede decision-making and, specifically, at the point when data is collected for processing. Within this debate, there have been other scholars that have tried to derive a right to an *ex-post* explanation from the right of access established in Art. 15 GDPR, under

which data subjects are granted a right to be informed about the existence of automated decision-making and to obtain meaningful information about the significance, envisaged consequences, and logic involved. Specifically, the data subject should be informed about the existence, purposes, and logic of data processing, and the intentions and legal consequences of such processing. By having this information, the data subject should be able to examine the lawfulness of data processing and invoke legal remedies. However, it is reasonable to doubt that the right of access grants a right to *ex-post* explanations of specific decisions already reached. Indeed, considering the semantics of Art. 15(1)(h), the phrase 'envisaged consequences' is future-oriented, suggesting that the data controller must inform the data subject of possible consequences of the automated decision-making before such processing occurs. This interpretation follows the timeline constraints of identical provisions in Art. 13(2)(f) and 14(2)(g) discussed above, which only allow for *ex ante* explanations. Data controllers are required to predict the possible consequences of their automated decision-making methods. The term 'envisaged' limits these predictions to *ex ante* explanations of system functionality, for instance concerning the general purpose of the system, or the type of impact to be expected from the type of decision it makes[200].

Within this debate, WP29 in its Guidelines stressed the importance of the 'right to an explanation' in relation to the broader principle of transparency, arguing that data subjects 'will only be able to challenge a decision or express their view if they understand how it has been made and on what basis'. These Guidelines also clarified that the information transmitted to the data subject should be sufficiently comprehensive to 'understand the reasons for the decision' but that this does 'not necessarily [require] a complex explanation

[200] S. Wachter – B. Mittelstadt – L. Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation (December 28, 2016)*, International Data Privacy Law, 2017.

of the algorithms used or disclosure of the full algorithm'. In this understanding, the 'right to an explanation' under the GDPR resembles the duty to give reasons. However, due to its non-binding nature, it is unclear to what extent the WP29 guidance can influence judicial interpretations[201].

3.1.3 Reconciling AI and GDPR

It has been argued that the GDPR would be incompatible with the use of AI, given that the GDPR is based on principles – purpose limitation, data minimization, the special processing of 'sensitive data', the limitation on automated decisions – that are incompatible with the technical functioning of the AI *per se*.

Nevertheless, this section shows that it is actually likely that the GDPR is interpreted in such a way as to reconcile both desiderata: protecting data subjects and enabling useful applications of AI. It is true that the full deployment of the power of AI requires collecting vast quantities of data concerning individuals and their social relations, and that it also requires processing of such data for purposes that were not fully determined at the time the data were collected. However, there are ways to understand and apply the data protection principles that are consistent with the beneficial uses of AI.

For instance, in order to respect the requirement of consent to be specific and purpose limitation, it should be linked to a flexible application of the idea of compatibility, that allows for the reuse of personal data when this is not incompatible with the purpose for which the data was collected. Moreover, reuse for statistical purposes is assumed to be compatible, and thus would in general be admissible (unless it involves unacceptable risks for the data subject). Even the principle of data-minimization can be understood in such a way as to enable a beneficial application of AI. This may involve in some context reducing the

[201] M. Fink – M. Finck, *Reasoned A(I)administration: explanation requirements in EU law and the automation of public administration*, European Law Review, 47, 3, 2022, p. 376 – 392.

‘personality’ of the data, namely the ease with which they can be connected to the individuals concerned, with measures such as pseudonymization, rather than focusing on the amount of personal data to be preserved. Furthermore, the information requirements established by the GDPR can also be met with regard to AI-based processing, even though the complexity of AI systems represents a difficult challenge. The information concerning AI-based applications should enable the data subjects to understand the purpose of the processing and its limits, without going into technical details. In addition, it should be recalled that the GDPR does not exclude automated decision-making, as it provides for ample exceptions – contract, law, or consent – to the general prohibition set forth in Art. 22(1). Lasty, the GDPR provisions on preventive measures, and in particular those concerning privacy by design and by default, should also not hinder the development of AI applications, if correctly designed and implemented, although they may entail some additional costs.

However, the aim of reconciling the GDPR and AI does not mean that such a balance can be found by referring to the GDPR alone. The GDPR rules need to be interpreted and consistently implemented, and appropriate guidance needs to be provided on concrete implication of the GDPR for particular processing activities. In various cases, the interpretation of undefined GDPR standards requires balancing competing interests: it requires determination of whether a certain processing activity, and the measures adopted are justified on balance, *i.e.*, whether the controller’s interests in processing the data and in (not) adopting certain measures are outweighed by the data subjects’ interests in not being subject to the processing or in being protected by additional or stricter measures. These assessments depend on both (a) uncertain normative judgements on the comparative importance of the impacts on the interests at stake and (b) uncertain forecasts concerning potential future risks. In the case of AI applications, the uncertainties

involved in applying indeterminate concepts and balancing competing interests are aggravated by the novelty of the technologies, their complexities, the broad scope of their individual and social effects. It is true that the principles of risk-prevention and accountability potentially direct the processing of personal data toward being a 'positive sum' game (where the advantages of the processing, when constrained by appropriate risk-mitigation measures, outweigh its possible disadvantages), and enable experimentation and learning, avoiding the over- and under-inclusiveness issues involved in the applications of strict rules. On the other hand, by requiring controllers to apply these principles, the GDPR offloads the task of establishing how to manage risk and find optimal solutions onto controllers, a task which may be both challenging and costly. The stiff penalties for non-compliance, when combined with the uncertainty as to what is required for compliance, may constitute a novel risk, which, rather than incentivizing the adoption of adequate compliance measure, may prevent small companies from engaging in new ventures.

No easy solution is available in the hyper-complex and rapidly evolving domain of AI technologies: rules may fail to enable opportunities and counter risks, but the private implementation of open standard, in the absence of adequate legal guidance, may also be unsatisfactory: '[Giving] appropriate content to the law often requires effort, whether in analyzing a problem, resolving value conflicts, or acquiring empirical knowledge […] [I]ndividuals contemplating behavior that may be subject to the law will find it more costly to comply with standards, because it generally is more difficult to predict the outcome of a future inquiry (by the adjudicator, into the law's content) than to examine the result of a past inquiry. They must either spend more to be guided properly or act without as much guidance as under rules'[202].

[202] L. Kaplow, *Rule vs standards: An economical analysis*, Duke Law Journal, 1992, p. 557–629.

Thus, the way in which the GDPR affects successful applications of AI in Europe also depends on what guidance data protection authorities and jurisprudential bodies – and more generally the legal system – are able to provide to controllers and data subjects. This would diminish the cost of legal uncertainty and would direct public actors to efficient and data protection-compliant solutions. Appropriate mechanisms may need to be devised, such as an obligation to notify data protection authorities when new applications based on profiling are introduced, but also the possibility of asking for preventive, non-binding, indications on whether and how such applications should be developed, and with what safeguards[203].

For these reasons, in the next sections some case law on the application of the GDPR in the context of the AI used by the public actor is presented.

3.2 The attempt of the jurisprudence to answer to the open questions

As shown in the previous sections, legal issues exist in relation to the use of these AI systems by the public actor that stems from the innate characteristic of the algorithms as they do not work according to a causal and deterministic logic, rather they develop machine learning paths which yield unpredictable results. As such, the system, constantly and automatically evolving, is not able to indicate the logic pathway which it has followed to reach the final result.

Thus, the necessity for the public actor to evaluate the compatibility of AI systems arises both considering the principles of constitutional and administrative law, with particular reference to those relating to due process, as well as the GDPR, given that most of these systems involve the processing of personal data.

[203] Panel for the Future of Science and Technology EPRS, *The impact of the General Data Protection Regulation (GDPR) on Artificial intelligence*, European Parliamentary Research Service Scientific Foresight Unit (STOA), (June 2020).

In this context, it is interesting to analyze relevant case laws as to understand the direction the jurisprudence has taken in solving legal issues related to the use of AI by the public actor. Particularly, considering the copious number of sentences held, the Italian administrative jurisprudence is examined in the following sections (3.2.1-3.2.2), as it has tried to give to the Italian public administration some directions by identifying the legal requirements, that in addition to Art. 22 of the GDPR, AI systems need to conform to so as to be used in procedures or administrative activities. Moreover, the C-333/22[204] decision of the CJEU is presented as it has imposed significant limits on the creation and use of algorithms and other forms of modern technology for security purposes by the public actor (section 3.2.3). Lastly, a judgment of a Duch Court is analyzed as it perfectly shows the balances between the interests of the public actor and citizens fundamental rights (section 3.2.4). Indeed, due to the lack of a clear regulatory framework in the field of automated public actor's decisions, and also due to the difficulty of the national legislators in following the continuous technological innovations, in recent years administrative jurisprudence has defined the principles to regulate public decision-making processes[205]. Rather than being an exhaustive analysis of the position of all Member States' jurisprudence, the following sections try to point out possible answers to the development and use of AI in the public field that has been given by the judges.

3.2.1 The initial orientation of the Italian administrative jurisprudence

The initial orientation of the Italian administrative judges (*i.e.*, Tar Lazio sec. III-*bis*, 10 October 2018, n. 9224, Tar Lazio, sec. III-*bis*, 10 September 2019, n. 10964) seems to

[204] CJEU, *Ligue des droits humains ASBL, BA v. L'organe de contrôle de l'information policière*, C-333/2022, (2022).
[205] C. Fusco, *The Use of Artificial Intelligence in the Decision-Making Processes of the Public Administration: Regulations and Executive Practice - The Case of the Italian Public Administration*, in (eds.) D. Marino - M. A. Monaca, *Innovations and Economic and Social Changes due to Artificial Intelligence: The State of the Art, Studies in Systems, Decision and Control*, Berlin, 2023.

exclude the admissibility of algorithms – and even more of AI systems – as sole instruments of the decision processing, considering the principle of 'instrumentality of the use of information technology in administrative procedures'[206]. Particularly, the Italian administrative judges (who dealt with a case relating to teacher transfers and assignment to public schools based on the results of an algorithm[207]), while recognizing the extendibility of the notion of 'administrative act' to the algorithm in order to give it full legal value and, thus, admitting its review by the judge, they deemed the use of just automated decisions with incisive effects on citizens without human intervention to be illegal. They considered the automated decisions incompatible with the principles of constitutional and administrative law, due to the automatism and opacity of the algorithm used[208]. Particularly, the ruling turned out to be very clear in excluding the possibility of a fully automated administrative activity as it was considered illegitimate to devolve to an impersonal algorithm the carrying out of the entire administrative procedure with the effect of totally replacing human activity[209].

This jurisprudential orientation shows the initial closure to the use of new technologies in the Italian administrative procedure, since the figure of the 'person in charge of the procedure' was prioritized, as aimed at giving a human face to the administration, as well as at avoiding dynamics of bureaucratic depersonalization. Accordingly, the use of digital solutions in the administrative process has to have a merely servant and instrumental function with respect to the activity of the public official in charge[210,211].

---

[206] Tar Lazio, sec. III-*bis*, 10 September 2019, n. 10964.
[207] Tar Lazio, sec. III-*bis*, 10 October 2018, n. 9224.
[208] C. Fusco, *The Use of Artificial Intelligence in the Decision-Making Processes of the Public Administration: Regulations and Executive Practice - The Case of the Italian Public Administration*, in (eds.) D. Marino - M. A. Monaca, *Innovations and Economic and Social Changes due to Artificial Intelligence: The State of the Art, Studies in Systems, Decision and Control*, Berlin, 2023.
[209] E. Carloni, *I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo*, Diritto amministrativo: rivista trimestrale, 2, 2020, p. 273-304.
[210] Tar Lazio, sec. III-*bis*, 10 September 2019, n. 10964.
[211] E. Carloni, *I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo*, Diritto amministrativo: rivista trimestrale, 2, 2020, p. 273-304.

3.2.2 A change of direction of the Italian jurisprudence

However, a change regarding the jurisprudential orientation on the use of AI within the administration of public functions can be observed. On this merit, through a series of decisions issued from 2019[212], the Consiglio di Stato[213] admitted the use of algorithms, albeit, initially, exclusively in the context of proceedings of a binding nature, and in particular in the context of repetitive proceedings, lacking in discretion, of a 'massive' or serial nature, such as those characterized by a simple automatic classification of numerous applications, in order to speed up the processes and ensure the accuracy and predictability of outcomes[214]. Moreover, moving on, the same judges expanded this indication, admitting algorithmic decisions also in the context of administrative proceedings characterized by discretion, taking care of identifying the minimum guarantees that the public actor is required to follow in these cases. Therefore, by admitting the possibility of using algorithms and, more broadly, AI systems, judges have highlighted the need for these systems to respect certain guarantees, derived from both domestic law principles and European law legislation[215].

This new direction emerges in few cases[216,217], where the Consiglio di Stato claims that the public administration must be able to exploit the potential of the so-called digital

[212] *Ex multis*, Consiglio di Stato, Sec. VI, 8 April 2019, n. 2270; Consiglio di Stato, Sec. VI, 13 Decembre 2019, n. 8472, Consiglio di Stato, Sec. VI, 13 December 2019, n. 8473, Consiglio di Stato, Sec. VI, 13 December 2019, n. 8474, Consiglio di Stato, Sec. VI, 04 February 2020, n. 881, Consiglio di Stato, Sec. VI, 04 June 2021, n. 1206.

[213] Please consider that in Italy the Consiglio di Stato represents the second-level administrative court.

[214] C. Fusco, *The Use of Artificial Intelligence in the Decision-Making Processes of the Public Administration: Regulations and Executive Practice - The Case of the Italian Public Administration*, in (eds.) D. Marino - M. A. Monaca, *Innovations and Economic and Social Changes due to Artificial Intelligence: The State of the Art, Studies in Systems, Decision and Control*, Berlin, 2023, p. 273-304.

[215] Consiglio di Stato, Sec. VI, 13 Decembre 2019, n. 8472.

[216] Consiglio di Stato, sec. VI, 8 April 2019, n. 2270; and Consiglio di Stato, Sec. VI, 13 Decembre 2019, n. 8472.

[217] Some teachers criticized their assignment far from their place of residence: the recruitment procedure, which had started with an application to be presented exclusively electronically on a platform set up by the Ministry (MIUR), responsible for data collection and processing, had already been dealt with by previous sentences held by the Tar. Access to the assignment criteria and to the source code of the software was

revolution and underlines the promised benefits in terms of efficiency and neutrality[218] also in public services. However, according to the Court, the coordinates of legitimacy of administrative action has to be given also in order to assure the accountability of the decision of the legal body that holds the power, which must be able to verify the logic and legitimacy of the choice and of the results. Moreover, since full knowledge implies transparency, the procedure, the decision, priorities assigned, and relevant data used in the automated decision processing must be made understandable externally. In this sense, industrial secrecy and privacy of companies are not relevant when they put their tools at the service of public administrations.

In addition, the Consiglio di Stato also refers to the GDPR. Indeed, Art. 13 and 14 of the GDPR require that the information of an automated decision-making process is given to the data subject and that in the case of a fully automated process significant information must be provided on the logic used, as well as the importance and expected consequences of this processing for the data subject. Moreover, Art. 15 of the GDPR grants the right to receive information relating to the existence of any automated decision-making processes, even if it has already started and even if it has produced a decision. Finally, Art. 22 of the GDPR affirms the right not to be subjected to automated decisions without human involvement that produce legal effects or similarly affect the individual.

requested, pursuant to art. 22 of Law 241 of 1990. The MIUR initially refused and then described the operation of the algorithm, commissioned to a company, without however providing technical data, as protected as an intellectual work and in any case not subject to access, as it cannot be qualified as administrative document. The TAR has ordered MIUR to issue copies of the software source codes to the applicants. The software code is considered a representation of activities of public interest, and it is not enough to indicate the operating criteria, also because it was in doubt that it did not respect the collective agreement of the school staff. It has been noted that, more than the transparency of the software, it would have been necessary to display its design criteria, leaving the administration with the task of verifying whether they had been respected.

[218] Consiglio di Stato, Sec. VI, 13 Decembre 2019, n. 8472.

Ultimately, in light of all these constitutional and legislative principles, the Consiglio di Stato indicates three cornerstone requirements of the use of automated decision processing by the public actor[219]:

(i) transparency: the algorithmic logic underlying the decision must always be motivated.

(ii) non-exclusivity: human involvement must be ensured, with the corollary of strict legality of fully automated decision processing.

(iii) non-discrimination: the quality of the data used as input for the decision must always be guaranteed, and above all, a reasonable and proportional assessment must be made of the use of automated decision processing.

Along this path, while admitting the use of the algorithm within the public decision-making, the administrative judge ultimately seeks to raise the bar of algorithmic legality by defining a set of conditions which makes the algorithmic decisions admissible[220]. Particularly, transparency lies at the heart of algorithmic due process constituting direct and specific application of Art. 42 of the Charter. Moreover, the identification of its centrality is strictly interconnected with the European discipline of the GDPR, whose recognizes transparency as a general principle, that applies also to algorithmic administrative procedures. Furthermore, it is interesting to note how in underlining the necessity of 'humanity' in the administration procedure, the most recent jurisprudence takes as a reference, precisely, Art. 22 of the GDPR (*i.e.*, a reference whose centrality had been highlighted before). This is because the GDPR states different principles directed to protect fundamental rights, which adopts an 'elastic' approach, inclined to admit, albeit conditioning it, the entry of automation into the heart of public decision-making processing. Lastly, the statement of

[219] Consiglio di Stato, Sec. VI, 13 Decembre 2019, n. 8472.
[220] E. Carloni, *I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo*, Diritto amministrativo: rivista trimestrale, 2, 2020, p. 273-304.

the principle of non-discrimination is of uttermost importance also for the corollaries it brings. First, the algorithmic rule used in the context of administrative decisions must in any case comply with the principles of impartiality and nondiscrimination, as well as with reasonableness and proportionality. Second, the need to base decision-making processes on 'correct' data calls into question the reliability and quality of the information on which the algorithms rely on. More specifically, this theme recalls the assessment of the quality of public data as a corollary to the principle of algorithmic non-discrimination. In these terms, the challenge within the AI-driven public actor paradigm is linked to the question of the quality of the data that is collected and processed as the prerequisite for making a decision. Therefore, this shifts the duty of the public actor to an organizational dimension, imposing it to design the algorithm in a way it respects specific rules and requirements. In this perspective, the quality of data stands as a fundamental moment of assurance of the algorithmic decision.

As shown, in these last decisions the Italian administrative judge, recognizing the benefits the AI systems could bring to the exercise of public functions, has shown openness to the introduction of new technologies. Nevertheless, being aware at the same time of the complexities of the new decision-making forms[221] and of the possible impacts on fundamental rights, the Court pointed out three cornerstone requirements of the use automated decision processing by the public actor. These are currently the major guidelines being followed in the public sector. However, gaps exist on the set of procedural rules and guarantees that should be guaranteed to citizens in order to ensure the protection of fundamental rights.

[221] *Ibid*.

3.2.3 Ligue des Droits Humains Case

On June 21st, 2022, the CJEU released its judgment regarding the compatibility of the EU Directive on Passenger Name Record Data[222] (PNR Directive) with the rights to privacy and personal data protection. *Ligue des droits humains* is a landmark decision, where the Court had the opportunity, among other aspects, to provide comprehensive guidelines on how large-scale predictive policing should take place. In so doing, the Court followed in the footsteps of La Quadrature du Net[223], and Opinion 1/15[224], which tackled a relentlessly growing kit of big-data-based security instruments.

On the merits of the case, the CJEU imposed significant limits on the creation and use of algorithms and other forms of modern technology for security purposes. *Ligue des droits humains* gave the Court the opportunity to do so because the PNR Directive obliges designated national security authorities, so-called Passenger Information Units (PIUs), to automatically process PNR data by comparison, not only against pre-existing databases, but also using 'pre-determined criteria'. The latter are algorithms which contain 'search criteria, based on the past and ongoing criminal investigations and intelligence, which allow to filter out passengers which corresponds to certain abstract profiles […]'[225]. Accordingly, pre-determined criteria serve to 'identify persons involved in criminal or terrorist activities who are, as of yet, not known to the law enforcement authorities'[226].

More specifically, the judgment restricted the 'use of artificial intelligence technology in self-learning systems (machine learning)', by prohibiting systems that are 'capable of modifying without human intervention or review the assessment process and, in particular,

[222] Directive 2016/681 - Use of passenger name record (PNR) data for the prevention, detection, investigation and prosecution of terrorist offences and serious crime.
[223] CJEU, *La Quadrature du Net u.a.*, C-511/18, (2020).
[224] CJEU, Opinion 1/15 of 26 July 2017.
[225] Commission Staff, *Accompanying the document report from the Commission to the European Parliament and the Council On the review of Directive 2016/681 on the use of passenger name record (PNR) data for the prevention, detection, investigation and proesecution of terrorist offences and serious crime*, (July 2020), p. 11.
[226] *Ibid*, p. 14.

the assessment criteria on which the result of the application of that process is based as well as the weighting of those criteria'[227]. Interestingly, the Court insists on the necessity of meaningful human intervention in predictive policing systems – a general principle already enshrined in Art. 11 of the Law Enforcement Directive. Central for its assessments was the 'opacity which characterizes the way in which artificial intelligence technology works', which may 'deprive the data subjects […] of their right to an effective judicial remedy'[228].

Moreover, the Court highlighted the risk of discrimination. While the PNR Directive already acknowledged such risks in its Art. 6(4), the Court also emphasized that this provision covers both direct and indirect discrimination[229]. This is crucial because pre-established criteria may be based on seemingly innocuous personal data, which may, however, be proxies of prohibited characteristics. For example, a person's address may also be used as a proxy for religion, race or ethnic origin. Algorithms must be 'targeted, proportionate and specific'[230] and thus non-discriminatory – a finding with wider implications in the context of migration, where non-discrimination must be (but is not) embedded in the discretionary decision-making process.

Furthermore, according to the Court, used technologies have to comply with a set of additional quality standards in order to incorporate 'incriminating' as well as 'exonerating circumstances'[231], thus bolstering their reliability and reducing false-positive rates. The Court stated that high false-positive rates, which are present in Member States' statistics, may undermine a system's suitability and proportionality[232]. Thus, the CJEU stressed the

[227] CJEU, *Ligue des droits humains ASBL, BA v. L'organe de contrôle de l'information policière*, C-333/2022, (2022), para. 194.
[228] *Ibid*, para. 195.
[229] *Ibid*, para. 197.
[230] *Ibid*, para. 198.
[231] *Ibid*, para. 200.
[232] *Ibid*, para. 123.

necessity of regular reviews of the pre-determined criteria's strict necessity[233]. Whilst acknowledging 'the fairly substantial number of 'false positives'', the CJEU stressed that 'the appropriateness of the system […] essentially depends on the proper functioning of the subsequent verification of the results […] by non-automated means'[234]. For that purpose, the Court determined that Member States must 'lay down clear and precise rules capable of providing guidance' for the review[235], which is meant to prevent both discriminatory results, and false matches to be transferred to the competent authorities, thus subjecting passengers to false suspicions of being involved in terrorist offences or serious crimes.

In addition, the Court underscored that the Data Protection Officer and national supervisory authorities must be equipped with robust rights of access to the content of pre-determined criteria[236].

It also aimed at reliable documentation and self-monitoring[237], as well as guaranteeing uniform administrative practices across PIUs in different Member States that observes the principle of non-discrimination. The results of individual human reviews must take preference over those of automated processing[238].

Finally, the judgment bolstered the right to an effective judicial remedy, as enshrined in Art. 47 of the Charter, stating that: 'The competent authorities must ensure that the person concerned […] is able to understand how those criteria and those programs work', so that they can 'decide with full knowledge of the relevant facts whether or not to exercise [their] right to the judicial redress', pursuant to Art. 13 of the PNR Directive[239]. This seems to

[233] *Ibid*, para. 201.
[234] *Ibid*, para. 124.
[235] *Ibid*, para. 205.
[236] *Ibid*, para. 212.
[237] *Ibid*, para. 207.
[238] *Ibid*, para. 208.
[239] *Ibid*, para. 216.

imply notification requirements in cases of verified positive matches which currently neither the PNR Directive nor most national transposition laws expressly contain[240].

3.2.3.1 Further considerations

Whereas the Court established an abundance of procedural safeguards to reign in the potential excesses of automated predictive threat detection, *Ligue des droits humains* also left a lot of open questions.

First, while the Court rightly flagged false positives as a potential hurdle to a system's proportionality, it did not clarify at what point a system just produces too many of them, thus rendering the system disproportionate. Although this point was raised in the oral hearing, the Court also never addressed the base rate fallacy undergoing the PNR system. The PNR Directive seeks to identify a very small number of potential terrorists and serious offenders within the general population of hundreds of millions of annual flight passengers. It – therefore like some other predictive policing systems – compels security authorities to look for the proverbial needle in the haystack. This can result in systemic flaws which make extremely high false positive rates a mathematical near-certainty. It remains to be seen whether mere procedural safeguards can succeed in saving the PNR system's suitability as long as its underlying base rate fallacy remains unaddressed.

Second, it remains unclear what purpose and form human interventions in the PNR system have to take, in particular, how humans are supposed to meaningfully engage with the PNR system's automated outputs[241]. The Court delegated the formulation of 'clear and precise rules' for human review to Member States[242] without providing them with much

[240] C. Thönnes – N. Vavoula, *Automated predictive threat detection after Ligue des Droits Humains: Implications for ETIAS and CSAM (Part I)*, VerfBlog, 2023.
[241] *Ibid*.
[242] CJEU, *Ligue des droits humains ASBL, BA v. L'organe de contrôle de l'information policière*, C-333/2022, (2022), para. 205.

guidance. That PIU officials will be capable of meaningfully engaging with the PNR system's automated outputs seems doubtful when, for the foreseeable future, they will be confronted with thousands of false matches.

Third, whereas the Court's insistence on substantive human review is extremely important to prevent automation bias, there also remain questions regarding how human review is supposed to prevent direct and indirect discrimination, hamstrung by a phenomenon known as 'selective adherence bias'. Recent studies[243] suggest that the 'human in the loop' may be predisposed to agree with those results of the automated processing that are more aligned with their personal pre-existing biases. Such biases may be based on socially induced stereotypes, beliefs and social identities, and result in selective adoption of algorithmic advice. It is known that humans tend to be susceptible to confirmation bias, meaning that they assign greater weight to information congruent with prior beliefs and less to content that contradicts them. Whereas automation bias may be more easily detected, it may be more difficult to detect and prevent selective adherence bias, especially in cases where the competent authorities share the same (*e.g.*, regional) biases as the PIU. This holds true especially when PIUs and competent security authorities receive an excess of potentially false matches. In practice, this could result in high risks of false suspicion for members of negatively stereotyped minority groups.

These gripes notwithstanding, the ruling does provide important guidelines on assessing other security-related instruments. Some of the aforementioned standards may be tailored to the PNR context. However, they pertain to features and risks that the PNR system

[243] S. Alon-Barkat – M. Busuioc, *Human-AI Interactions in Public Sector Decision-Making: "Automation Bias" and "Selective Adherence" to Algorithmic Advice*, Journal of Public Administration Research and Theory, 2022.

shares with other security instruments designed at preventively detecting threats in large data pools through automated processing[244].

3.2.4 The Dutch SyRI Case

In addition to the orientation of the Italian jurisprudence, another case of national courts *NCJM et al. and FNV vs. The State of the Netherlands*[245] is presented considering the novelty of the analysis and findings by the Hague District Court. Indeed, this is one of the first judgments addressing human rights implications of the use of AI in the public sector and states' respective obligations to ensure transparency of AI processes[246]. Particularly, the case challenged the Dutch government's use of System Risk Indication (SyRI) – an algorithm designed to identify potential social welfare fraud[247]. It was brought by a coalition of Dutch civil society organizations and two Dutch citizens challenging the Dutch government's use of SyRI. The plaintiffs submitted that the legislation governing SyRI and the algorithm's use violated the right to private life and personal data protection, as provided under Art. 8 of the ECHR and Art. 7 and 8 of Charter. They also claimed a violation of Art. 6 and 13 of the ECHR, as well as of Art. 5, 6, 13, 14, 22 and 28 of GDPR. In conclusion, the Court ruled that neither the legislation governing SyRI nor its use met

[244] C. Thönnes – N. Vavoula, *Automated predictive threat detection after Ligue des Droits Humains: Implications for ETIAS and CSAM (Part I)*, VerfBlog, 2023.
[245] District Court of the Hague, 6 March 2020, n. 865.
[246] A. Rachovitsa - N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.
[247] SyRI is a big-data analysis system that ran under the auspices of the Dutch Ministry of Social Affairs and Employment. The aim of the system is to prevent and combat fraud in the areas of income-dependent schemes, taxes and social security. It could be used upon request by one of the so-called 'cooperation associations', namely governmental bodies and certain Dutch municipalities. Following the linkage of many siloed datasets held by government agencies, the aggregated data were fed into the SyRI algorithm. The algorithm's risk model used several unknown risk indicators (for example, related to taxes, health insurance, residence, education), on the basis of which it detected increased risk of irregularities by generating risk profiles of cases suspected of presenting a higher likelihood of fraud. The submission of such a risk report could result in further investigation by relevant authorities. Risk notifications were included in a register for 2 years. Individuals were not informed when a risk report had been created for them and were not able to gain any insights into how a decision was reached. SyRI's legal basis lay in two main domestic laws: Section 65 of the Work and Income (Implementation Organisation Structure) Act (SUWI Act) taken together with Chapter 5a of the Decree concerning the rules for tackling fraud by exchanging data and the effective use of data known within the government with the use of SyRI (SUWI Decree).

the requirements laid down in Art. 8(2) of the ECHR for an interference with the exercise of the right to private life to be necessary and proportionate[248].

3.2.4.1 The main questions before the Court

The main questions before the Court were whether the SyRI legislation and the use of SyRI (i) constituted an interference with Art. 8 of the ECHR; (ii) pursued a legitimate aim; (iii) had a basis in law; and (iv) were necessary and proportionate restrictions of the right to private life[249].

In order to answer these questions (i), the Court paid particular attention to the extent and seriousness interference as a factor that weighs heavily in the assessment of the necessity and proportionality of a restriction of the right to privacy. The Court had to clarify two issues in this regard: first, the nature of SyRI and, second, the legal effect of a risk report. Starting with the nature of SyRI, the Court held that it was 'unable to assess [...] the precise nature of SyRI because the State has not disclosed the risk model and the indicators of which the risk model is composed or may be composed'[250]. During the court proceedings, the state did not provide the Court with objectively verifiable information arguing that a disclosure could lead to citizens adjusting their conduct in order to avoid detection of fraud. The Court maintained that this was a deliberate choice on behalf of the state, which was also reflected in the absence of transparency in the SyRI legislation as to how the system's decision model functioned. The Court went on to find that, 'the SyRI legislation does not provide room for unstructured data collection with the use of SyRI'[251]. Although the amount of data that could be used was substantial, the data categories were

[248] A. Rachovitsa - N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.
[249] *Ibid*.
[250] District Court of the Hague, 6 March 2020, n. 865.
[251] *Ibid*.

exhaustively enumerated. The Court accepted that, as currently implemented, SyRI did not entail any use of deep learning and data mining. Nonetheless, links between data sets were established, which, in turn, led to results suggestive of an increased risk of committing fraud. In addition to this, the law not only failed to preclude the use of predictive analyses, deep learning and data mining, but also expressly allowed for the adjustment of a risk model and the development of models with new indicators. Therefore, in its view, the application of SyRI was aligned with deep learning and self-learning systems[252]. As regards the risk profiles, it was not possible for the Court to ascertain whether these were being developed, although it was found that risk profiles based on existing factual data were intrinsic to SyRI's application. The Court emphasized that the 'SyRI legislation does not provide for a duty of disclosure to those whose data is processed in SyRI'[253]. Similarly, the law did not provide for an obligation to notify data subjects individually when a risk report had been submitted. Data subjects were not informed unless an investigation had been initiated in response to a risk report, which did not happen as a matter of course. The Court did not find it satisfactory that the only statutory obligation was to announce the start of a SyRI project by way of a publication in the Government Gazette or that access to the register of risk reports was only granted upon request after the data processing had taken place. The second matter to be elucidated in assessing the intrusiveness of the interference was the legal effect of a risk report. According to the Court, the submission of a risk report based on the application of SyRI constituted profiling within the meaning of Art. 4(4) of the GDPR. Thus, it suggested that, although the use of SyRI in and of itself was not intended to have legal effect, 'a risk report does have a similarly significant effect on the private life of the person to whom the risk report pertains'[254]. A risk report could be

[252] A. Rachovitsa - N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.
[253] District Court of the Hague, 6 March 2020, n. 865.
[254] *Ibid*.

stored for 2 years and could be used by participants in the SyRI project for a period of 20 months. Other investigative authorities could also be notified of the report upon request. Even if a risk report never resulted in an investigation or sanctions, the effect on the private life of the data subject would remain pronounced. This effect, in conjunction with the data subject's inability to be reasonably aware of the processing of their data, led the Court to conclude that the interference with the right to private life was extensive and serious[255].

Following question (ii), the Court proceeded to examine whether the SyRI legislation pursued a legitimate aim. The matter was not disputed as combating fraud in social security and welfare is considered a legitimate and important goal. Therefore, the Court admitted that the legislation was pursuing a legitimate aim as it was provided by the law.

The question (iii) that the Court had to answer was whether the interference with the right to private life had a sufficiently accessible and foreseeable legal basis. Accessibility and foreseeability are the so-called quality of law requirements and refer to the need that a piece of legislation must be formulated with sufficient precision so as to enable individuals to regulate their conduct. In other words, individuals must be able to access and foresee, to a degree that is reasonable in the circumstances, the consequences which a given action may entail. Interestingly, the accessibility and foreseeability criteria regain their relevance in the context of rapid technological developments. This is because the domestic legislator struggles to sufficiently regulate such developments and may be intentionally obscure or vague in doing so. More specifically, in the SyRI case, following the ECtHR's approach in *S. and Marper v. United Kingdom*[256], the Court chose to 'leav[e] undiscussed in its review whether the SyRI legislation is sufficiently accessible and foreseeable and as such affords an adequate legal basis'[257], as required under Art. 8(2)

[255] A. Rachovitsa - N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.
[256] ECtHR, *S. and Marper v. The United Kingdom*, App. 30562/04 and 30566/04, (2008).
[257] District Court of the Hague, 6 March 2020, n. 865.

of the ECHR. This choice was justified on the basis 'that the SyRI legislation in any case contains insufficient safeguards for the conclusion that it is necessary in a democratic society' and an assessment of the adequacy of the legal basis was thus not made[258]. Given the fact that the Court had reservations on whether the legislation governing SyRI met the accessibility and foreseeability criteria, it is unfortunate that it drew no formal conclusion on this matter[259].

The (iv) and final question addressed by the Court was whether the use of SyRI and the legislation pertaining to it were necessary and proportionate restrictions on the right to private life. In order to resolve this question, the Court relied heavily upon the EU law principles of transparency, purpose limitation and data minimization. It maintained that the legislation was insufficiently transparent and verifiable and that the use of SyRI entailed an interference with the right to respect for private life, which was unnecessary and disproportionate to the purpose of combating fraud. This conclusion was grounded on three findings. First, the Court found that neither the legislation nor the use of SyRI respected the principle of transparency. The principle is grounded in Art. 8(2) of the Charter and Art. 5(1)(a) and 12–15 of the GDPR. The normative scope of the principle encompasses a right for data subjects to access their data and obligations imposed upon data controllers to inform data subjects of data collection and processing activities. Crucially, transparency also serves an enabling function for the effective enjoyment of other data subject rights[260]. In the case of the SyRI legislation, it provided no information on the objective factual data that could justify an inference of increased risk in individual cases. Nor was there any clear insight into the functioning of the risk model, such as the

[258] *Ibid*.
[259] A. Rachovitsa - N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.
[260] R. Polcák, *Article 12. Transparent Information, Communication and Modalities for the Exercise of the Rights of the Data Subject,* in (eds.) C. Kuner – L. A. Bygrave – C. Docksey, The EU General Data Protection Regulation (GDPR) A Commentary, 2020, p. 401–402.

algorithms or risk-analysis method used[261]. The Court thus questioned 'how a data subject could be able to defend themselves against the fact that a risk report has been submitted about him or her'[262] or 'be aware that their data were processed on correct grounds'[263]. Second, the importance of transparency in connection with the ability to verify the risk model and risk indicators was all the greater since the use of the risk model entailed the risk of discriminatory effects – unintentional or otherwise. On this matter, the Court clarified that due to the large amounts of data that qualify for processing and the use of risk profiles 'there is in fact a risk that SyRI inadvertently creates links based on bias, such as a lower socio-economic status or an immigration background'[264]. No evidence was presented to suggest that safeguards were put in place to neutralize the risk of discriminatory effects. Third, the law contained insufficient safeguards relating to the principles of purpose limitation and data minimization. Although the legislation provided an exhaustive list of data categories to be used in SyRI projects, the personal data eligible for processing in SyRI were essentially limitless. For this reason, the assessments conducted by national authorities on whether a given interference with private life was necessary and proportionate in light of the specific purpose of a SyRI project were critical. However, the Court found that there was no legal requirement for either a review by an independent third party prior to data processing approval or a comprehensive review of the necessity of using SyRI in specific instances. Considering the lack of respect for the principle of transparency in conjunction with inadequate safeguards, the Court, thus, found a violation of Art. 8 of the ECHR and declared that the SyRI legislation 'ha[s] no binding effect with

[261] District Court of the Hague, 6 March 2020, n. 865.
[262] *Ibid.*
[263] *Ibid*.
[264] *Ibid*.

respect to [the admissible claimants] and on the individuals whose interests these parties promote'[265].

In the end, the state did not appeal the judgment, which therefore is now final[266].

3.2.4.2 Important remarks on the SyRI case

Analyzing the Court's decision, a prominent and overarching theme underpinning all aspects of the SyRI case is how the lack of transparency, first, inhibited data subjects from effectively exercising their rights and, second, undermined the exercise of the judicial function.

Indeed, algorithmic transparency, in this instance, is not merely complementary to the human rights law claims but provides for strong grounds for finding a violation of the right to privacy. This comes to confirm how AI systems challenge the traditional concepts of transparency and accountability and, at the same time, reinforce the need for 'radical transparency about the impact of an AI system in the information environment'[267]. More specifically, the Court found that the legislation provided little, if any, insight into the risk model and risk indicators used, the objective factual data that could justifiably lead to the inference of an increased risk or the data processed in SyRI projects. Crucial information concerning the algorithm's use was deliberately kept secret creating opacity[268]. In this regard, the Netherlands' refusal to disclose additional information, on the grounds that citizens would otherwise 'game the system', is an argument invoked by many countries with regard to different uses of algorithmic systems in different public areas[269]. Although

---

[265] *Ibid*.
[266] A. Rachovitsa - N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.
[267] Secretary General, *Report of the Special Rapporteur on the promotion and protection of the rights to freedom of opinion and expression*, (August 18th, 2017).
[268] J. Niklas, *Human Rights-Based Approach to AI and Algorithms Concerning Welfare Technologies*, in (eds.) W. Barfield, *The Cambridge Handbook of the Law of Algorithms*, 2021, p. 517.
[269] AlgorithmWatch, *Automating Society Report 2020*, (October 2020).

there is merit in protecting the public interest in investigating and prosecuting crimes, there is equally significant merit, to say the least, in protecting the public interest in the ability of those concerned to obtain information about how these systems function and how their rights are affected. In this sense, the models are not merely a matter of internal concern to welfare bureaucracies but a matter of public interest. It is arguable that the same standard of rule of law concerning the publicness and transparency of law also needs to be applied to algorithmic systems used by public authorities, so that citizens know what is expected of them[270].

Furthermore, the absence of information on SyRI hampered the Court's ability to pass judgment on many crucial points and, therefore, to exercise proper judicial oversight of the application of the law in accordance with international human rights. On multiple occasions, the Court expressly held that it was unable to answer legal questions such as the legal nature of SyRI, whether risk profiles were developed in the course of its use or whether the risk of discrimination was sufficiently neutralized[271]. The question then arises as to how the lack of information is to be appreciated when assessing a potential human rights violation. Accordingly, when the lack of transparency and the absence of safeguards are affecting the rights of individuals, the state's scant justification have led the Court to conclude that neither the legal basis nor the use of SyRI met the requirements of Art. 8 of the ECHR. Therefore, in this instance, the state's failure to provide a convincing explanation for the lack of transparency and alternative safeguards to protect data subjects' rights was critical for affirming a violation of the right to privacy[272].

[270] A. Rachovitsa - N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.
[271] District Court of the Hague, 6 March 2020, n. 865.
[272] A. Rachovitsa - N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.

Another important finding in this case was that the state had employed hidden algorithmic risk models by specifically and exclusively targeting neighborhoods inhabited by low-income and minority background demographics. Predictive analytics, algorithms and other forms of AI are highly likely to reproduce and exacerbate biases reflected in existing data and policies. Identifying and counteracting such biases in designing the digital welfare state is important and requires precisely what was found to be lacking in the Dutch legislation: transparency in law and in practice about how an AI system works and broad-based inputs into policymaking processes[273]. The public, and especially those affected by the welfare system, need to be able to understand and evaluate the processes and outcomes buried deep within the algorithms[274]. Those targeted by algorithmic systems are the least likely to be able to defend themselves against intrusions and the ensuing negative consequences. On this matter, the Court ruled that the SyRI legislation contained no safeguards to neutralize the risk of discriminatory and stereotyping effects. The right to social security encompasses the right to access and maintain benefits, without discrimination – whether in law or in fact, direct or indirect – on grounds such as race, color, sex, age, language or national or social origin, especially with regard to individuals belonging to disadvantaged and marginalized groups. Curiously, none of these claims were raised by the plaintiffs and, even if they had been raised, it would have been difficult for the Court to entertain them precisely due to the lack of information and evidence. This may have been the reason that the plaintiffs did not raise the complaints in the first place. Proving potential and/or indirect discrimination is a challenging task in any context – all the more so when algorithmic systems and prediction analytics come into play. Without any information on the functioning of the algorithm, it is impossible for plaintiffs to

[273] Secretary General, *Report of the Special Rapporteur on the promotion and protection of the rights to freedom of opinion and expression*, (August 18th, 2017).
[274] *Ibid*.

successfully argue for (a risk of) discrimination without insight into the risk factors used by the algorithm. The difficulties of substantiating the (potential) discriminatory effect of algorithmic systems cast doubt on whether data protection rules and international human rights law, as they currently stand, are well-suited to address risks posed to specific groups. The same concern also applies to conceptualizing new types of discriminatory harm and/or societal harm, some of which we may not even be able to anticipate at this point in time. In cases where no specific individual is necessarily discriminated against, but rather groups, people and neighborhoods are being targeted, anti-discrimination laws and human rights provisions arguably fall short of grasping and articulating the issues at stake.

In conclusion, these types of challenges bring into play not only existing limitations of substantive human rights law (for example, the protective scope of the principle of non-discrimination or other human rights) but also crucial questions around admissibility (for instance, the inability to claim a violation of group rights under current human rights law and, therefore, lack of standing) or the evidence and burden of proof required to substantiate such claims[275].

3.2.5 Issues left open: the limit of the case in civil law orders

Despite the jurisprudential attempts to provide specific requirements for the public actor while using AI technologies in public decision processing, gaps exist on the set of procedural rules and guarantees that should be granted to citizens. Moreover, the issue related to ‘depersonalization’ in the relationship between administration and citizen requires a re-calibration of algorithmic criteria by placing the individual at the center of the decision-making logic. Additionally, while principles of transparency and non-

[275] A. Rachovitsa - N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.

discrimination seem to be more easily applicable to automated decision-making processes based on mechanical and deterministic algorithms, on the contrary, it does not appear to be sustainable in the case of the use of algorithms such as probabilistic or machine learning algorithms, characterized by a level of opacity that makes them unintelligible[276]. In sum, these issues show the need to define the conditions for a full compatibility between the constitutional principles and the development of increasingly more sophisticated AI systems[277].

As seen in the previous sections, in the absence of specific legislative provisions, it is the administrative judge who has so far played a central role in the development of reflection, and in the definition of legal coordinates, with respect to the topic of public actor's algorithmic decisions. A judge who is called upon to confront algorithmic decisions and who is now emerging as a precursor in the attempt to provide the first useful answers to govern these phenomena. However, the path toward building a 'technological due process' within the paradigm of the AI-driven public actor still appears to be on a preliminary stage.

Therefore, there is no doubt that the considerations outlined above make the need clear for an explicatory intervention by the legislator[278]. In this context, it is thus necessary to analyze the provisions of the AI Act.

[276] C. Fusco, *The Use of Artificial Intelligence in the Decision-Making Processes of the Public Administration: Regulations and Executive Practice - The Case of the Italian Public Administration*, in (eds.) D. Marino – M. A. Monaca, *Innovations and Economic and Social Changes due to Artificial Intelligence: The State of the Art, Studies in Systems, Decision and Control*, Berlin, 2023.

[277] E. Carloni, *I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo*, Diritto amministrativo: rivista trimestrale, 2, 2020.

[278] C. Fusco, *The Use of Artificial Intelligence in the Decision-Making Processes of the Public Administration: Regulations and Executive Practice - The Case of the Italian Public Administration*, in (eds.) D. Marino – M. A. Monaca, *Innovations and Economic and Social Changes due to Artificial Intelligence: The State of the Art, Studies in Systems, Decision and Control*, Berlin, 2023.

3.3 The AI Act: the landing point?

As mentioned in the previous sections, on August 2nd, 2024, the AI Act entered into force. As stated in Recital 1:

> 'The purpose of this Regulation is to improve the functioning of the internal market by laying down a uniform legal framework in particular for the development, the placing on the market, the putting into service and the use of artificial intelligence systems (AI systems) in the Union, in accordance with Union values, to promote the uptake of human centric and trustworthy artificial intelligence (AI) while ensuring a high level of protection of health, safety, fundamental rights as enshrined in the Charter of Fundamental Rights of the European Union (the 'Charter'), including democracy, the rule of law and environmental protection, to protect against the harmful effects of AI systems in the Union, and to support innovation. This Regulation ensures the free movement, cross-border, of AI-based goods and services, thus preventing Member States from imposing restrictions on the development, marketing and use of AI systems, unless explicitly authorized by this Regulation'.

Moreover, under Recital 6:

> 'Given the major impact that AI can have on society and the need to build trust, it is vital for AI and its regulatory framework to be developed in accordance with Union values as enshrined in Article 2 of the Treaty on European Union (TEU), the fundamental rights and freedoms enshrined in the Treaties and, pursuant to Article 6 TEU, the Charter. As a prerequisite, AI should be a human-centric technology. It should serve as a tool for people, with the ultimate aim of increasing human well-being'.

In light of these recitals, the EU states that it wants to build a human centric AI aimed at guaranteeing citizens' fundamental rights. Therefore, the new legal framework puts technology in a perspective of servility to humans being, as it should not overstep the

boundaries outlined within the Charter. In this sense, the AI Act can be considered to be the landing point of the AI regulation, which in the absence of previous specific provisions required the use of ancillary legislation such as the GDPR or, likewise, the intervention of judicial authority.

In light of this, although the debate on the clear application of the AI Act is still in its starting phase, given that the text is going to be applied, for a large part, from 2026, it is interesting to analyze its most important provisions in order to understand whether (i) the objectives stated by the European legislator in the recitals have actually been achieved; (ii) in the context of the AI-driven public actor, the AI Act is sufficient to protect citizens from the changing balances within this new paradigm.

3.3.1 Scope of application of the AI Act

The legal framework applies to both public and private actors inside and outside the EU as long as the AI system is placed on the Union market, or its use has an impact on people located in the EU. The obligations affect providers[279], deployers[280], authorized representatives[281], importers[282], as well as distributors[283] of AI systems.

There are certain exceptions to the application of the AI Act. Research, development and prototyping activities that take place before an AI system is released on the market are not subject to its provisions. Additionally, AI systems that are exclusively designed for

[279] Provider means a natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system or a general-purpose AI model developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge.

[280] Deployer means a natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity.

[281] Authorized representative means a natural or legal person located or established in the Union who has received and accepted a written mandate from a provider of an AI system or a general-purpose AI model to, respectively, perform and carry out on its behalf the obligations and procedures established by this Regulation.

[282] Importer means a natural or legal person located or established in the Union that places on the market an AI system that bears the name or trademark of a natural or legal person established in a third country.

[283] Distributor means a natural or legal person in the supply chain, other than the provider or the importer, that makes an AI system available on the Union market.

military, defense or national security purposes are also exempt, regardless of the type of entity carrying out those activities.

The AI Act provides for a different set of obligations depending on the inherent risk associated with the AI system/practice that is being used[284]. According to the so-called risk-based approach[285], the AI Act lays down: (i) prohibitions on certain AI practices (unacceptable risk/highest risk level); (ii) specific requirements for high-risk AI systems and obligations imposed on operators of such systems; (iii) harmonized transparency rules for certain AI systems which are outside the scope of those identified as 'high risk' and are not deployed for a prohibited practice; (iv) horizontal obligations for all GPAI models. In addition, the AI Act lays down rules on market monitoring, market surveillance, governance, and measures in support of innovation.

For the purpose of the AI Act, all those involved in using the aforementioned systems – whether as a provider, user, distributor, importer – will be subject to a level of regulatory scrutiny. Particularly, the rules established by the AI Act apply to providers placing on the market or putting into service AI systems in a non-discriminatory manner, namely irrespective of whether the said providers are physically present or established within the EU or in a third country, and to users of AI systems who are physically present or established within the EU. Moreover, it applies to providers and users of AI systems who are physically present or established in a third country, to the extent the output produced

[284] Please consider that during the negotiation of the AI Act, the co-legislators added the category of 'general purpose AI models'. However, this new category is inconsistent with a truly risk-based approach.

[285] The AI Act is claimed to follow a risk-based approach – one that tailors the choice and design of regulatory instruments based on the level of risk, according to the rule: the higher the risk, the stricter the rules. On this merit, Recital 26 of the AI Act points out that: 'In order to introduce a proportionate and effective set of binding rules for AI systems, a clearly defined risk-based approach should be followed. That approach should tailor the type and content of such rules to the intensity and scope of the risks that AI systems can generate. It is therefore necessary to prohibit certain unacceptable AI practices, to lay down requirements for high-risk AI systems and obligations for the relevant operators, and to lay down transparency obligations for certain AI systems'. Hence, the underlying objective of the AI Act's risk-based approach is to strike an optimal (or proportionate) balance between innovation and the benefits of AI systems on the one hand, and the protection of fundamental values such as safety, health and fundamental rights on the other.

by those systems is used in the EU. Furthermore, the AI Act applies to importers and distributors of AI systems, to product manufacturers placing on the market or putting into service an AI system together with their product and under their own name or trademark, as well as to authorized representatives of providers, which are established in the EU[286].

3.3.2 Prohibited artificial intelligence practices and high-risk systems

The AI Act lays down a ban on a limited set of uses of AI. Particularly, chapter II of the Regulation covers all those AI systems whose use is considered unacceptable as contradicting EU values, especially by violating fundamental rights. As the Commission put it, those AI systems could 'provide novel and powerful tools for manipulative, exploitative and social control practices' which are 'particularly harmful and abusive and should be prohibited because they contradict Union values of respect for human dignity, freedom, equality, democracy and the rule of law and fundamental rights enshrined in the Charter, including the right to non-discrimination, to data protection and to privacy and the rights of the child'[287]. Particularly, the prohibition covers the place on the market, putting into service or use of certain AI systems intended to distort human behavior, whereby physical or psychological harm is likely to occur. These intrusive practices include AI used for adverse behavioral influencing, social scoring, and large-scale surveillance[288]. More

[286] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.

[287] Recital 28 of the AI Act.

[288] According to Art. 5 of the AI Act the following artificial intelligence systems should be prohibited: (a) AI system that deploys subliminal techniques beyond a person's consciousness or purposefully manipulative or deceptive techniques, with the objective, or the effect of materially distorting the behaviour of a person or a group of persons by appreciably impairing their ability to make an informed decision, thereby causing them to take a decision that they would not have otherwise taken in a manner that causes or is reasonably likely to cause that person, another person or group of persons significant harm; (b) AI system that exploits any of the vulnerabilities of a natural person or a specific group of persons due to their age, disability or a specific social or economic situation, with the objective, or the effect, of materially distorting the behaviour of that person or a person belonging to that group in a manner that causes or is reasonably likely to cause that person or another person significant harm; (c) AI systems for the evaluation or classification of natural persons or groups of persons over a certain period of time based on their social behaviour or known, inferred

specifically, Art. 5 of the AI Act explicitly bans harmful AI practices that are considered to be a threat to people's safety and rights, because of the 'unacceptable risk' they create[289]. Therefore, starting from February 2nd, 2025, the ban of these systems will be applicable. In addition, chapter III of the AI Act creates a separate tier of high-risk AI systems and contains specific rules for those AI systems that cause a high risk to health and safety or fundamental rights of natural persons. In line with a risk-based approach, those high-risk AI systems are permitted on the European market on a restricted basis, with specific controls in place to support safe use and subject to compliance with certain mandatory requirements and an ex-ante conformity assessment. According to Recital 47, 'AI systems could have an adverse impact on the health and safety of persons, in particular when such systems operate as safety components of products [...]. It is important that the safety risks that may be generated by a product as a whole due to its digital components, including AI

---

or predicted personal or personality characteristics, with the social score leading to either or both of the following: (i) detrimental or unfavourable treatment of certain natural persons or groups of persons in social contexts that are unrelated to the contexts in which the data was originally generated or collected; (ii) detrimental or unfavourable treatment of certain natural persons or groups of persons that is unjustified or disproportionate to their social behaviour or its gravity; (d) AI system for making risk assessments of natural persons in order to assess or predict the risk of a natural person committing a criminal offence, based solely on the profiling of a natural person or on assessing their personality traits and characteristics; this prohibition shall not apply to AI systems used to support the human assessment of the involvement of a person in a criminal activity, which is already based on objective and verifiable facts directly linked to a criminal activity; (e) AI systems that create or expand facial recognition databases through the untargeted scraping of facial images from the internet or CCTV footage; (f) AI systems to infer emotions of a natural person in the areas of workplace and education institutions, except where the use of the AI system is intended to be put in place or into the market for medical or safety reasons; (g) AI systems for the use of biometric categorisation systems that categorise individually natural persons based on their biometric data to deduce or infer their race, political opinions, trade union membership, religious or philosophical beliefs, sex life or sexual orientation; this prohibition does not cover any labelling or filtering of lawfully acquired biometric datasets, such as images, based on biometric data or categorizing of biometric data in the area of law enforcement; (h) the use of 'real-time' remote biometric identification systems in publicly accessible spaces for the purposes of law enforcement, unless and in so far as such use is strictly necessary for one of the following objectives: (i) the targeted search for specific victims of abduction, trafficking in human beings or sexual exploitation of human beings, as well as the search for missing persons; (ii) the prevention of a specific, substantial and imminent threat to the life or physical safety of natural persons or a genuine and present or genuine and foreseeable threat of a terrorist attack; (iii) the localisation or identification of a person suspected of having committed a criminal offence, for the purpose of conducting a criminal investigation or prosecution or executing a criminal penalty for offences referred to in Annex II and punishable in the Member State concerned by a custodial sentence or a detention order for a maximum period of at least four years.

[289] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.

systems, are duly prevented and mitigated'. 'AI systems identified as high-risk should be limited to those that have a significant harmful impact on the health, safety and fundamental rights of persons in the Union and such limitation minimizes any potential restriction to international trade, if any'[290].

Stand-alone AI systems in eight areas, with mainly health, safety, and fundamental rights implications, which are explicitly listed in Annex III of the AI Act, are also considered high-risk. According to the Commission, 'as regards stand-alone AI systems, meaning high-risk AI systems other than those that are safety components of products, or which are themselves products, it is appropriate to classify them as high-risk if, in the light of their intended purpose, they pose a high risk of harm to the health and safety or the fundamental rights of persons, taking into account both the severity of the possible harm and its probability of occurrence and they are used in a number of specifically pre-defined areas specified in the Regulation. The identification of those systems is based on the same methodology and criteria envisaged also for any future amendments of the list of high-risk AI systems'[291]. The list of (stand-alone) high-risk AI systems in Annex III contains a limited number of AI systems whose risks have already materialized or are likely to materialize in the near future.

This list covers a wide range of applications including AI systems deployed in relation to: (i) biometric systems; (ii) critical infrastructure (*e.g.*, road traffic and the supply of water, gas, heating and electricity) and protection of the environment; (iii) education and vocational training; (iv) employment, workers management and access to self-employment; (v) access to and enjoyment of essential private services and public services and benefits; (vi) law enforcement; (vii) migration, asylum and border control management; (viii) administration of justice and democratic process.

[290] Recital 46 of the AI Act.
[291] Recital 52 of the AI Act.

Considering the possible harm to citizens' fundamental rights, under the AI Act, high-risk AI systems shall comply with specific mandatory requirements set out in Section 2 of Chapter 3, which are described in the following sections (3.3.2.1 – 3.3.2.7). When ensuring compliance with those requirements, the intended purpose of the use of the high-risk AI system and the risk management system to be established by the provider should be taken into account.

In this context, it should be stressed that Art. 111(2) of the AI Act excludes (*i.e.*, the AI Act will not apply to) high-risk AI systems already placed on the market or put into service before application of the AI Act, unless after that date the AI system undergoes significant changes in their design or intended purpose[292]. However, this exception does not apply to the public actor, as it is clarified that 'in any case, the providers and deployers of high-risk AI systems intended to be used by public authorities shall take the necessary steps to comply with the requirements and obligations of this Regulation by 2 August 2030'.

3.3.2.1 Risk Management System

A risk management system must be established, implemented, documented, and maintained as part of an overall quality management system. The risk management system must comprise 'a continuous iterative process run throughout the entire lifecycle of a high-risk AI system, requiring regular systematic updating'[293] and include the following steps: (i) identification and examination of the known and foreseeable risks associated with the high-risk AI system; (ii) estimation and assessment of the risks that may arise when the high-risk AI system is used in accordance with its intended purpose and under

[292] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.
[293] Art. 9 of the AI Act.

conditions of reasonably foreseeable misuse; (iii) evaluation of other possibly arising risks based on the analysis of data gathered from the post-market monitoring system referred to in Art. 72 of the AI Act; (iv) adoption of suitable risk management measures, which should take into account the generally acknowledged state of the art (as reflected, for instance, in relevant harmonized standards or common specifications) and ensure that the overall residual risk (which must be communicated to the user) of the high-risk AI systems is judged acceptable, provided that the said high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse. Particularly, in order to ensure the integration of appropriate safeguards and risk management measures, the following should be achieved: (i) elimination or reduction of risks; (ii) adequate mitigation and control measures in case risks cannot be eliminated; (iii) training and provision of concise, adequate and comprehensible information to users as regards the risks that may emerge when the high-risk AI system under examination is used in accordance with its intended purpose and under conditions of reasonably foreseeable misuse. The measures taken to eliminate or reduce risks related to the use of the high-risk AI system must give due consideration to technical knowledge, experience, education, training to be expected by the user and the environment in which the system is intended to be used. Finally, the AI Act states that, when implementing the risk management system (and adopting suitable risk management measures), specific consideration must be given to whether the high-risk AI system is likely to be accessed by or have an impact on children[294].

[294] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.

3.3.2.2 Data Governance: use of high-quality data sets

The use of high-quality data is essential for the performance of AI systems, particularly when techniques involving the training of models are used, with a view to ensuring that the high-risk AI systems perform safely and as intended and that they do not become the source of discrimination. It is recalled that the White Paper on AI[295] emphasized that 'without data, there is no AI'. The functioning of many AI systems, and the actions and decisions to which they may lead, depend on the data set on which the systems have been trained. Therefore, the AI Act introduces specific measures in relation to the high-risk AI applications in order to ensure that where it comes to the data used to train AI systems, the EU's values and rules are respected, specifically in relation to safety and existing legislative rules for the protection of fundamental rights. Thus, the aim is to safeguard that those datasets are representative of diverse populations, as well as to safeguard that when used to train AI systems they are unlikely to lead to prejudicial decisions. The production of data should engage all parts of society, *i.e.* all relevant stakeholders that may be affected by AI systems should be consulted and involved in those systems' development and implementation ('stakeholder participation'), and that more diverse datasets should be created in order to fairly reflect the societies and communities which AI systems are increasingly affecting. Those measures also need to consider biases embedded in the algorithms themselves. For instance, developers set the parameters for machine learning algorithms, and the choices they make will intrinsically reflect the developers' own beliefs, preferences, and prejudices. Therefore, the researchers' and developers' AI teams should also be diverse and representative of wider society.

[295] European Commission, *White Paper on Artificial Intelligence - A European approach to excellence and trust*, (February 19th, 2020).

To this end, the AI Act stresses that high-risk AI systems which make use of techniques involving the training of models with data should be developed on the basis of training, validation and testing data sets that comply with specific quality criteria[296]. Particularly, in order to ensure high quality training, validation and testing data sets, the implementation of appropriate data governance and management practices is required (such as the requirement to check the suitability or unbiasedness of the data and to guarantee that there are no gaps or shortcomings in those datasets). Furthermore, training, validation and testing data sets should be sufficiently relevant, representative, free of errors and complete in view of the intended purpose of the system. They should also have the appropriate statistical properties, including as regards the persons or groups of persons on which the high-risk AI system is intended to be used. In order to protect the right of others from the discrimination that might result from the bias in AI systems, the providers of high-risk AI systems should be able (subject to appropriate safeguards such as pseudonymization or encryption) to process also special categories of personal data, as a matter of substantial public interest, in order to ensure the bias monitoring, detection and correction in relation to high-risk AI systems. Furthermore, appropriate data governance and management practices should apply to the development of high-risk AI systems other than those which make use of techniques involving the training of models[297].

[296] Art. 10 of the AI Act.
[297] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.

3.3.2.3 Technical Documentation

In order to provide information on how high-risk AI systems have been developed and how they perform throughout their lifecycle, and facilitate supervision and enforcement, the availability of appropriate technical documentation is essential.

Particularly, the establishment of appropriate technical documentation is necessary to verify compliance of high-risk AI systems with the requirements set out in the AI Act and, at the same time, to allow national competent authorities and notified bodies to assess the compliance of the AI systems with those requirements. More specifically, the technical documentation should be drawn up before the high-risk AI system is placed on the market or put into service, should be kept up-to date[298], and should include, at a minimum, specific elements such as the general characteristics, capabilities and limitations of the system, algorithms, data, training, testing and validation processes used as well as documentation on the relevant risk management system.

The technical documentation – which is set out in Annex IV of the AI Act – should provide at least the following information: (i) a general description of the AI system; (ii) a detailed description of the elements of the AI system and of the process for its development; (iii) detailed information about the monitoring, functioning and control of the AI system; (iv) a description of the appropriateness of the performance metrics for the specific AI system; (v) a detailed description of the risk management system; (vi) a description of any change made to the system through its lifecycle; (vii) a list of the harmonized standards applied and, if no such harmonized standards have been applied, a detailed description of the solutions and a list of other relevant standards and technical specifications applied; (viii) copy of the EU declaration of conformity; (ix) a detailed description of the system in place

[298] Art. 11 of the AI Act.

to evaluate the AI system performance in the post-market phase (including the post-market monitoring plan referred to in Art. 72(3) of the AI Act)[299,300].

3.3.2.4 Record-keeping in order to ensure traceability and accountability

Considering the potential impact of the use of AI systems on citizens' lives, according to the AI Act, it is necessary to ensure their traceability[301].

Thus, the data sets, the decisions made by the systems and the processes made by the AI systems' decisions should be logged and well documented. This leads to an increase in transparency and enables the identification of the reasons why an AI-decision was erroneous, which in turn facilitates auditability and explainability of the algorithmic decision-making process. Furthermore, the requirement of accountability responds to the necessity to explain and to hold to account the decision-making processes of an AI system. If an AI system causes harm, it has to be possible to provide an adequate explanation, which in turn has to be auditable by a competent authority. In addition, it is imperative to identify who can be considered responsible – and held liable – in the event of failure of an AI system[302]. As such, for high-risk AI systems, the logging capabilities shall provide, at a minimum: (i) recording of the period of each use of the system (start date and time and end date and time of each use); (ii) the reference database against which input data has been checked by the system; (iii) the input data for which the search has

[299] Annex IV of the AI Act.
[300] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.
[301] Art. 12 of the AI Act.
[302] N. T. Nikolinakos, The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament, in (eds.) N. T. Nikolinakos, EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series, Berlin, 2023, p. 327-741.

led to a match; (iv) the identification of the natural persons involved in the verification of the results.

3.3.2.5 Transparency and provision of information

In order to tackle the opacity issue, the AI Act underlines that high-risk AI systems should be designed and developed in such a way to ensure that their operation is sufficiently transparent so that users are able to interpret the system's output and use it appropriately[303]. A certain degree of transparency should be required with a view to achieving compliance with the obligations of providers and users of high-risk AI systems, as set out in Chapter 3, Section 2 of the AI Act. Particularly, high-risk AI systems should be accompanied by relevant documentation and instructions of use in an appropriate digital format or otherwise that include concise, complete, correct, and clear information that is relevant, accessible, and easily understandable by users. This information should comprise, *inter alia*: (i) the identity and the contact details of the provider, and, where applicable, of its authorized representative; (ii) the characteristics, capabilities and limitations of the performance of the high-risk AI system; (iii) the changes to the high-risk AI system and its performance which have been pre-determined by the provider at the moment of the initial conformity assessment; (iv) the human oversight measures put in place to facilitate the interpretation of the outputs of AI systems by the users; (v) the expected lifetime of the high-risk AI system and any necessary maintenance and care measures to ensure its proper functioning; (vi) a description of the mechanism included within the AI system that provides users with the opportunity to properly collect, store and interpret the logs, where relevant[304].

[303] Art. 13 of the AI Act.
[304] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial*

3.3.2.6 Human oversight

One of the questions that have been left open by the application of the GDPR is the necessity of human oversight during the automated decision processing. Following ample discussion on this theme, in the Annex to the European Parliament Resolution of 20 October 2020[305], the European Parliament stated that high-risk AI technologies should be human-centric, developed, deployed and used in a manner that guarantees full human oversight at any time and allows full human control to be regained when needed. Particularly, it states that 'decisions made or informed by artificial intelligence, robotics and related technologies should remain subject to meaningful human review, judgment, intervention and control. The technical and operational complexity of such technologies should never prevent their deployer or user from being able to, at the very least, trigger a fail-safe shutdown, alter or halt their operation, or revert to a previous state restoring safe functionalities in cases where the compliance with Union law and the ethical principles and legal obligations laid down in this Regulation is at risk'[306].

Based on these recommendations – and not only – the AI Act stresses the need to exercise enhanced oversight over specific AI systems. In particular, it states that high-risk AI systems should be designed and developed in such a way, including with appropriate human-machine interface tools, that natural persons can effectively oversee their functioning during the period in which the AI system is in use[307]. For this purpose, appropriate human oversight measures should be identified (and built into the system, when technically feasible) by the provider of the high-risk AI system before its placing on

Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series, Berlin, 2023, p. 327-741.

[305] Framework of ethical aspects of artificial intelligence, robotics and related technologies, *European Parliament resolution of 20 October 2020 with recommendations to the Commission on a framework of ethical aspects of artificial intelligence, robotics and related technologies*, (October 20th, 2020).

[306] *Ibid*.

[307] Art. 14 of the AI Act.

the market or putting into service, aiming at minimizing (or preventing, where possible) the risks to health, safety or fundamental rights and the implications and damaging effects that may emerge when a high-risk AI system is used in accordance with its intended purpose or under conditions of reasonably foreseeable misuse, especially when such risks persist. Such human oversight measures should ensure that the high-risk AI system is subject to in-built operational constraints that cannot be overridden by the system itself and is responsive to the human operator, and that the individuals to whom human oversight has been assigned have the necessary competence, training and authority to carry out that role. In particular, the assigned individuals should be enabled to: (i) fully understand the capacities and limitations of the high-risk AI system, and to duly monitor its operation, so that signs of dysfunctions and unexpected performance can be timely detected and addressed; (ii) remain aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system ('automation bias'); (iii) correctly interpret the high-risk AI system's output; (iv) decide when not to use the high-risk AI system or otherwise disregard, override or reverse the output of the high-risk AI system; (v) intervene on (or abort) the operation of the high-risk AI system or interrupt the system through a 'stop' button or a similar procedure[308].

As such, Art. 14 of the AI Act requires that high-risk AI systems shall be designed and developed in such a way, including with appropriate human machine interface tools, that they can be effectively overseen by natural persons during the period in which the AI system is in use. It has been noted that ensuring the presence of human oversight in the AI system is necessary not only to prevent and reduce risks – to health, safety or fundamental rights, as the standard states – that may result from 'improper' though

[308] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.

'reasonably foreseeable' use of the AI system, but also when that system is used 'in accordance with its intended purpose'. The provision thus considers human surveillance as an indispensable form of prevention even with regard to properly carried out data processing, because these too are likely to produce risks that only the presence of humans can minimize. Paragraphs 3 and 4 of Art. 14 go on to establish how human surveillance is ensured, *i.e.,* by introducing it and integrating it into the AI system before it is placed on the market, and what effects it is to produce, *i.e.,* full management of the operation of the AI system and control of its effects by humans. Indeed, it is necessary that the person to whom the supervisory tasks are assigned: (i) fully understands the capacities and limitations of the high-risk AI system; (ii) is able to duly monitor its operation and detect its malfunctions in a timely manner; (iii) is aware of the possible tendency of automatically relying or over-relying on the output produced by a high-risk AI system ('automation bias'); (iv) is able to correctly interpret the high-risk AI system's output; (v) is able to decide not to use it or to disregard, override or reverse its output, as well as to intervene on the operation of the system even by interrupting it through a 'stop' button or a similar procedure.

Therefore, from the AI Act can be inferred how human surveillance represents a fundamental safeguard in the implementation of such AI systems.

However, this safeguard remains in any case indeterminate in its implementation dimension. Indeed, with respect to its application at least three critical issues emerge. First of all, some interpretive problems arise with regard to the exact definition of 'human surveillance', as it is clear that human intervention can only be useful if the person in charge of the control is actually able to assess the adequacy of the indications on the basis of which the AI system works, that is the process that led to the final result[309].

[309] C. Fusco, *The Use of Artificial Intelligence in the Decision-Making Processes of the Public Administration: Regulations and Executive Practice - The Case of the Italian Public Administration*, in (eds.) D. Marino – M.

Moreover, the formulation of Art. 14 has been criticized due to it seems that some of its requirements are too stringent, unreasonable, unfeasible and impossible to implement in practice and should, therefore, be modified to exclude certain circumstances where human oversight is neither necessary nor appropriate. There are critical cases where human oversight is needed continuously. In some cases, human oversight can lead to delays, in others, accuracy of outputs could even be undermined by human interventions. For less critical situations, detailed human involvement may not be necessary or proportionate. For instance, human oversight is not considered to be feasible in certain (complex) autonomous systems, such as fully autonomous vehicles, factory production equipment, or cybersecurity systems where autonomous decision-making in real time is critical to the application. In those cases, the effectiveness of human supervision may be detrimental. Thus, the requirement to 'fully' understand the capacities and limitations of a high-risk system is effectively impossible since humans cannot maintain efficient oversight over a system that is much faster than themselves. Indeed, in those cases, maintaining humans in the loop defeats the purpose of automation as the efficiency of AI systems come to the fore when replicating tasks that previously required human cognitive abilities at scale.

Therefore, the requirements of Art. 14 of the AI Act should be determined based on the specifics of the individual use cases. Furthermore, it can be argued that the obligation to establish a system of human oversight differs from human intervention. Oversight is required to ensure that these systems are operating in the intended manner and without harming individuals. Oversight involves a pre- and post-system output review to evaluate performance and outcomes but may not involve human intervention while the AI system is functioning. Indeed, intervention would require human beings to take specific action at

A. Monaca, Innovations and Economic and Social Changes due to Artificial Intelligence: The State of the Art, Studies in Systems, Decision and Control, Berlin, 2023.

certain points in the process to either validate or reject an AI product or system's decision/recommendation/action, which can result in not being feasible in certain situations[310].

3.3.2.7 Accuracy, robustness and cybersecurity

Within Art. 15 of the AI Act also accuracy, robustness and cybersecurity requirements are provided. Particularly, the AI Act stresses that high-risk AI systems shall be designed and developed in such a way that they perform consistently throughout their operational lifecycle and meet, in the light of their intended purpose, an appropriate level of accuracy, robustness and cybersecurity in accordance with the generally acknowledged state of the art[311]. The level of accuracy and the relevant accuracy metrics of high-risk AI systems shall be correctly reflected during all life cycle phases, properly communicated to the end users and found in the accompanying instructions of use. High-risk AI systems should be resilient, adequately identifying the possible threats to the system and dealing with errors, faults or inconsistencies (design faults, technical faults, environmental threats) during all life cycle phases, especially due to the systems' interaction with natural persons or other systems. The robustness of high-risk AI systems may be achieved through the integration of safety and security-by-design mechanisms and the implementation of technical redundancy solutions and verification/validation methods ensuring that outcomes are reproducible, including backup or fail-safe plans. The increasingly autonomous high-risk AI systems using, *inter alia*, machine-learning and deep-learning techniques (*i.e.,* systems that continue to learn after being placed on the market or put into service) should ensure

[310] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.
[311] Art. 15 of the AI Act.

that possibly biased outputs due to outputs used as an input for future operations are properly tackled with the implementation of appropriate mitigation measures. Furthermore, high-risk AI systems should be resilient and protected against vulnerabilities (security gaps) that can be exploited by unauthorized third parties altering the use or performance of the systems for malicious purposes (*e.g.*, hacking/cyberattack). Appropriate technical measures need to be implemented to ensure cybersecurity of the high-risk AI systems, including testing (in order to identify and mitigate the risks of hacking and cyber-attacks) and establishing processes capable of assessing the safety risks involved. The technical solutions should include measures to prevent attacks targeting the data (data poisoning/manipulation of training data), inputs designed to cause the model to make a mistake ('adversarial examples'), model evasion (*i.e.*, classifying the data according to the attacker's will) and model inversion (*i.e.*, infer the model parameters). It should be stressed that, on the cybersecurity robustness criteria, the AI Act puts forward the presumption of compliance for high-risk AI systems when they are certified or when a statement of conformity has been issued under a cybersecurity scheme pursuant to the Cybersecurity Act[312,313] or the Cyber Resilience Act[314].

[312] Regulation (EU) 2019/881 of the European Parliament and of the Council of 17 April 2019 on ENISA (the European Union Agency for Cybersecurity) and on information and communications technology cybersecurity certification and repealing Regulation (EU) No 526/2013 (Cybersecurity Act).
[313] N. T. Nikolinakos, The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament, in (eds.) N. T. Nikolinakos, EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series, Berlin, 2023, p. 327-741.
[314] Regulation (EU) 2024/2847 of the European Parliament and of the Council of 23 October 2024 on horizontal cybersecurity requirements for products with digital elements and amending Regulations (EU) No 168/2013 and (EU) 2019/1020 and Directive (EU) 2020/1828 (Cyber Resilience Act).
Please note that Recital 51 of the Cyber Resilience Act states that 'Products with digital elements classified as high-risk AI systems pursuant to Article 6 of Regulation (EU) 2024/1689 of the European Parliament and of the Council which fall within the scope of this Regulation should comply with the essential cybersecurity requirements set out in this Regulation. Where those high-risk AI systems fulfil the essential cybersecurity requirements set out in this Regulation, they should be deemed to comply with the cybersecurity requirements set out in Article 15 of Regulation (EU) 2024/1689 in so far as those requirements are covered by the EU declaration of conformity or parts thereof issued under this Regulation'.

3.3.2.8 Fundamental rights impact assessment

The uniqueness of the forthcoming EU legal framework also lies in its intent to ensure a delicate balance between fostering the development of cutting-edge technologies and safeguarding safety and fundamental rights of individuals. Indeed, the AI Act introduces in the global AI regulatory landscape an innovative approach based on the risk introduced by the AI system for health, safety, fundamental rights, environment, and rule of law of natural persons. To this end, among the key requirements that the high-risk AI systems must comply with, the AI Act introduces a Fundamental Rights Impact Assessment (FRIA). The obligation to conduct FRIA under Art. 27 AI Act is in line with and complementary to other more limited and specific impact assessment obligations that public and private entities in the EU are already required to do under other legislations (*e.g.*, the GDPR and the Digital Services Act), adopting a focus on the variety of potential impacts on fundamental rights that high-risk AI systems may have[315].

The FRIA acts as a verification process suitable for understanding whether a violation of fundamental rights can be perpetrated in each specific use case. Particularly, ss further specified by Recital 96, the impact assessment should identify the deployer's relevant processes in which the high-risk AI system will be used in line with its intended purpose and should include a description of the period and frequency in which the system is intended to be used. By relying on the documentation shared by the provider, the deployer needs to identify the natural persons and the groups of people that might be affected by the processing. Moreover, deployers should also determine measures to be taken in the case of the realization and materialization of those risks. Furthermore, according to Recital 96, the deployer shall arrange a review process of handling such a potential infringement

[315] S. Bertaina – I. Biganzoli – R. Desiante – D. Fontanella – N. Inverardi – I. G. Penco – A. Cosentini, *Fundamental Rights and Artificial Intelligence Impact Assessment: A New Quantitative Methodology in the Upcoming Era of Ai Act*, 2024.

of rights. It is required to provide the natural persons with the safeguard of human oversight and ‘complaint handling and redress procedures, as they could be instrumental in mitigating risks to fundamental rights in concrete’. This process aims to empower individuals in the direct protection of their rights actively, thus requiring the design of this phase of compliance. Once the deployer performs the assessment, this should be reported to the market surveillance authority and updated if necessary. By imposing these stringent requirements, the EU is underscoring the importance of accountability and transparency in using AI technologies. Furthermore, the obligation to report these assessments to the market surveillance authority ensures that any identified risks are systematically addressed. This approach reflects a broader commitment to integrating ethical considerations into the development and deployment of AI, aligning with the EU’s wider agenda of promoting trustworthy AI[316].

However, there are some critical profiles that deserve further consideration. The first issue is the entity who should conduct the assessment. According to the text of the AI Act, the assessment will have to be conducted by the deployer before putting into service the system. However, this is a relevant limitation, as the deployer does not intervene in the design and development phases of the system, which are instead carried out by the provider. It would be important that the assessment could be conducted by the developer (maybe together with the deployer) and not by the deployer alone in the design and development phases. Indeed, it is the developer who has the possibility to intervene in the development stages of the algorithm and to opt for trustworthy by design. The same principle is not new to EU legislation, which has already applied it with great success in the GDPR with the concept of ‘privacy by design’. While it is true that it would be useful to

[316] G. De Gregorio – M. Fasciglione – F. Paolucci – O. Pollicino, *Compliance through Assessing Fundamental Rights: Insights at the Intersections of the European AI Act and the Corporate Sustainability Due Diligence Directive,* MediaLaws, 2024.

provide a first assessment conducted by the provider before the system is placed on the market, this does not exclude a further assessment that will be conducted during the lifecycle of the AI system by the deployer, who knows the context of use best of all. A second aspect of attention is that such an assessment is only required for high-risk systems. It is not possible to exclude *a priori* that there may be an infringement of fundamental rights also in areas not considered high risk by the AI Act. Moreover, despite Art. 27 states that 'The AI Office shall develop a template for a questionnaire, including through an automated tool, to facilitate deployers to implement the obligations of this Article in a simplified manner', it is yet not clear which method deployers will have to use to conduct the FRIA.

In any case, it should be taken into account that AI Act is evolutionary because it imposes specific obligations in terms of FRIA and consequently sanctions for non-compliant operators. The analysis and quantification of any potential detriment of fundamental rights will not be any more a voluntary action, a best practice, or a recommendation, and this is undoubtedly a great achievement in a civil technological society. However, several issues remain open, particularly regarding how such assessment will be conducted[317].

3.3.3 Transparency obligations for providers and deployers of certain AI systems and GPAI models

Following the risk-based approach, the AI Act provides for few obligations for certain AI systems that do not fall within the category of high-risk AI systems. As it is underlined in the Regulation, 'certain AI systems intended to interact with natural persons or to generate content may pose specific risks of impersonation or deception irrespective of whether they

[317] S. Bertaina – I. Biganzoli – R. Desiante – D. Fontanella – N. Inverardi – I. G. Penco – A. Cosentini, *Fundamental Rights and Artificial Intelligence Impact Assessment: A New Quantitative Methodology in the Upcoming Era of Ai Act*, 2024.

qualify as high-risk or not. In certain circumstances, the use of these systems should therefore be subject to specific transparency obligations without prejudice to the requirements and obligations for high-risk AI systems'[318]. The objective is to allow people to make informed choices or withdraw from a given situation[319]. That being considered, providers shall ensure that AI systems intended to directly interact with natural persons are designed and developed in such a way that the concerned natural persons are informed that they are interacting with an AI system, unless this is obvious from the point of view of a natural person who is reasonably well-informed, observant and circumspect, taking into account the circumstances and the context of use[320]. In this context, the AI Act sets specific transparency requirements for: (i) providers of AI systems, including General Purposes AI (GPAI) systems, generating synthetic audio, image, video or text content; (ii) deployers of an emotion recognition system or a biometric categorization system; (iii) deployers of an AI system that generates or manipulates image, audio or video content constituting a deep fake; (iv) deployers of an AI system that generates or manipulates text which is published with the purpose of informing the public on matters of public interest.

Moreover, specifically on general purpose AI systems, the AI Act introduces horizontal obligations for all GPAI models, which includes keeping up-to-date and making available, upon request, technical documentation to the AI Office and national competent authorities. It also includes providing certain information and documentation to downstream providers for the purpose of compliance with the AI Act. There are some additional requirements for models with systemic risks, which would include performing model evaluation, making risk assessments and taking risk mitigation measures, ensuring an adequate level of

[318] Recital 132 of the AI Act.
[319] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.
[320] Art. 50 of the AI Act.

cybersecurity protection, and reporting serious incidents to the AI Office and national competent authorities.

Under the AI Act, classification of GPAI models as presenting systemic risks depends on the capability, either based on a quantitative threshold of the cumulative amount of compute used for its training measured in floating point operations (FLOPs), or on an individual designation decision of the Commission that takes into account criteria listed in Annex XIII (*e.g.*, the number of parameters, quality and size of the dataset, input and output modalities or the reach measures in business users). Regarding copyright, the AI Act states that providers of GPAI models need to put in place a policy to respect Union copyright law, as well as make publicly available a sufficiently detailed summary about the content used for training of the general-purpose AI model, based on a template provided by the AI Office. Recital 107 explains that this summary should not be technically detailed but comprehensive at a general level, while taking into due account the need to protect trade secrets and confidential business information. Finally, Art. 53 includes rules to align GPAI to the risk-based approach of the AI Act. Providers of such systems will be obliged to provide all the information and elements to downstream providers of high-risk AI systems so that they can comply with the respective requirements, including for the purpose of conformity assessment.

3.3.4 Remedies

In addition to the above-mentioned provisions, the AI Act provides for an entire section of possible remedies to infringement of citizens' rights. Particularly, the AI Act establishes a right to lodge a complaint with a market surveillance authority, under Art. 85 (section 3.3.4.1) and a right to explanation of individual decision-making, under Art. 86 (section 3.3.4.2).

The provision of remedies is of the utmost importance as in order to offer effective judicial protection (enshrined on Art. 47 of the CFREU) being 'one of the most often used Charter right in legal proceedings', also covering 'decision taken with the support of AI technologies'[321]. In this sense, and after describing how '[u]sing AI can challenge the right to an effective remedy in different ways', particularly when parties involved have not the information and/or the knowledge to understand the systems' functioning, it is clear that '[w]ithout access to information, individuals may not be able to defend themselves, assign responsibility for the decisions affecting them, appeal any decision negatively affecting them or have a fair trial, which includes the principle of equality of arms and adversarial proceedings'[322,323].

However, it should be noted that the AI Act does not offer a right to a specific remedy, such as the right to compensation in Art. 82 of the GDPR. Nevertheless, it is worth noting individuals can rely on other regulations, such as the GDPR and liability laws, to address any harm caused by AI systems.

3.3.4.1 Lodging complaints directly with the relevant authorities

According to the AI Act, EU and national law already provides effective remedies to natural and legal persons whose rights and freedoms are adversely affected by the use of AI systems. Without prejudice to those remedies, any natural or legal person having grounds to consider that there has been an infringement of the provisions of this Regulation should be entitled to lodge a complaint to the relevant market surveillance authority[324]. Therefore, 'without prejudice to other administrative or judicial remedies, complaints to the relevant

[321] European Union Agency for fundamental rights, *Fundamental Rights Report*, (June 9th, 2020), p. 75.
[322] *Ibid.*, p. 76
[323] J. Covelo De Abreu, *The 'Artificial Intelligence Act' Proposal on European e-Justice Domains Through the Lens of User-Focused, User-Friendly and Effective Judicial Protection Principles*, in (eds.) H. S. Antunes – P. M. Freitas - A. L. Oliveira - C. Martins Pereira - E. Vaz De Sequeira – L. Barreto Xavier*, Multidisciplinary Perspectives on Artificial Intelligence and the Law*, Berlin, 2023, p. 397-414.
[324] Recital 170 of the AI Act.

market surveillance authority may be submitted by any natural or legal person having grounds to consider that there has been an infringement of the provisions of this Regulation'[325].

In this regard, it should be taken into account that that the provision of a right to lodge a complaint to the relevant market surveillance authority has been included only as a result of lengthy and complex negotiations, since the first version of the proposed AI Act made no mention of remedial mechanisms that could be activated directly by the users of AI systems. Therefore, in order to guarantee the right to an effective remedy under Art. 47 of the CFREU this introduction is very important.

On the merits of this right, it is interesting to note how the provision established for an unusually large personal scope of exercise, as there is practically no requirement of standing (*i.e.*, the only ground to be considered is that there has been an infringement of the provisions of the AI Act). This is clearly different from other instruments, such as the GDPR, where it is explicitly clarified that data subjects may submit a complaint only if the processing of personal data relates to them.

3.3.4.2 Right to an explanation

As mentioned before, the AI Act introduces a provision called right to explanation of individual decision-making, posing a new duty on the AI deployer as 'any affected person subject to a decision which is taken by the deployer on the basis of the output from an high-risk AI system listed in Annex III […] and which produces legal effects or similarly significantly affects him or her in a way that they consider to adversely impact their health, safety and fundamental rights shall have the right to request from the deployer clear and meaningful explanations on the role of the AI system in the decision-making procedure

[325] Art. 85 of the AI Act.

and the main elements of the decision taken'[326]. Such an explanation right includes: (i) the role of the AI system in the decision-making procedure; (ii) the main parameters of the decision taken; and (iii) the related input data.

With this provision, the AI Act seems to translate constitutional values (such as due process) in the context of automated decision-making, by fostering the principle of the rule of law. In a modern society, individuals are vulnerable and deserve further protection. Accordingly, the AI Act is requiring the explanation of the logic of deployers' actions and decisions, trying to guarantee a (re)balance of powers. In this scenario, the AI Act seems to take a position on the debate that arose around the existence of the right to explanation in the GDPR, which was arguably inspired by the right to a reasoned decision. Indeed, both these rights aim to provide the decision subjects the reasons behind the decision to potentially start a trial, assert the legality of the decision, defend their rights in the best possible conditions and decide how to behave with full knowledge of the relevant facts. Also, similarly, this right tries to rebalance the informational and power asymmetry between decision-makers and decision-subjects. Hence, arguably the right to explanation can be considered an AI-based declination of the right to a reasoned decision. However, there is a fundamental difference between the right to a reasoned decision and the right to explanation: whereas the reasoning must inform that the measure was adopted *secundum legem*, the explanation is sufficient when it explains that the decision was not taken *contra legem*. Thus, whereas the right to a reasoned decision must provide information concerning the subsumption of algorithmic parameters to rules of positive law or principles of law, the explanation does not have to provide this information because the right to explanation is released from the obligation to follow the parameters of

[326] Art. 86 of the AI Act.

administrative law and procedures[327]. Considering this scenario, it is interesting to try to understand if the right to a reasoned decision 'absorbs' the right to explanation. In this regard, considering the content requirements of such rights, it is arguable that the right to a reasoned decision does not fully absorb the right to explanation. Indeed, whereas the right to a reasoned decision has stringent requirements (provided for by case law) and must provide information concerning the subsumption of algorithmic parameters to rules of positive law or principles of law, the explanation does not have to provide this information because the right to explanation is 'released' from such obligation. Moreover, the AI Act add some further content requirements (the role of the AI system in the decision-making procedure, the parameters of the decision, and the related input data): this approach seems to tackle the problem of identifying the elements deemed to be essential to delegate the decision to an AI system and, at the same time, provide information which enables the decision-subject's understanding of the decision and eventually challenge it. In other words, the AI Act seems to identify the information which may help public authorities to comply with its reason-giving duties when deferring a recommendation provided by an AI system – even if the reason-giving requirements are more demanding (for a thorough and interesting analysis, see here). In this regard, it is arguable that the content requirements (to date) identified in the AI Act are not sufficient. Without dwelling here on what should be required in an explanation, arguably some elements can be fundamental to enable one's understanding of a decision, such as (i) input data having decisional impact or (ii) data features.

A further point should be stressed. The right to explanation in the AI Act is applicable only when decisions have significant impact, and therefore it has a more limited applicability

[327] J. Dirutigliano, *Some considerations on the relationship between the right to a reasoned decision and the right to explanation in the proposal of the artificial intelligence act*, The digital constitutionalist the future of constitutionalism, 2023.

than the right to a reasoned decision (also as a result of the latter being a general principle of EU law – its applicability is warranted in all contexts of individual decision-making in the EU). Not having to explain decisions that do not have a significant impact is, however, a reasonable approach taken by the EU since it may be necessary to strike a balance between the need for explanation and the costs associated with the difficulties of explaining automated decisions. The AI Act, in this sense, formalizes what such cases are. Indeed, the utility of explanations must be balanced against the cost of generating them. Accordingly, not every decision should be explained, but only those which significantly impact the decision-subjects. Especially, what renders a decision problematic and thus worthy of an explanation is the impact such a decision may have without knowing how and why it has been taken.

In light of these brief considerations, the introduction by the AI Act of an explanation duty addressed to private and public actors can be problematic. As noted by Hofmann, '[m]ixing public and private obligations is problematic since each have different legal obligations as to their procedures. Arguably, the use of AI in public decision-making should better be integrated into a general EU administrative procedures act and address specific effects of ADM on decision-making and rule-making procedures'[328]. In this regard, the application of the right to explanation should take inspiration from the right to a reasoned decision. Indeed, automated decision-making processes are governed by rules and transparencies policies specific to the sector in which they are used: the explanation's content should consider such rules, and also the context and goal for which the right to explanation is exercised. A 'general' right to explanation applicable to all areas of law risks not being effective because it would ignore the sector-specific features and the protection of other competing rights. Therefore, the transparency requirements and the information to provide

[328] H. C. H. Hofmann, *Automated Decision-Making (ADM)*, EU Public Law University of Luxembourg Law Research, 2023.

should differ depending on the type of right the explanation must safeguard and the purpose for which such right has been exercised[329].

3.3.5 Applicability of the AI Act within the AI-driven public actor paradigm

Considering the provisions analyzed above, it is necessary to examine whether the AI Act establishes specific provisions in case of use of AI by the public actor.

Preliminarily, it should be taken into account that the AI Act applies a risk-based approach to providers, deployers, importers, distributors and users regardless they are public or private entities.

That being said, considering that, depending on the circumstances regarding its specific application, use, and level of technological development, AI may generate risks and cause harm to public interests and fundamental rights that are protected by Union law, the AI Act classifies some AI systems that are used for public functioning as high-risk systems. Particularly, Annex III[330] identifies as high-risk AI systems the ones that allow the 'Access to and enjoyment of […] essential public services and benefits:

> (a) AI systems intended to be used by public authorities or on behalf of public authorities to evaluate the eligibility of natural persons for essential public assistance benefits and services, including healthcare services, as well as to grant, reduce, revoke, or reclaim such benefits and services;
>
> […]
>
> (c) AI systems intended to evaluate and classify emergency calls by natural persons or to be used to dispatch, or to establish priority in the dispatching of emergency first

[329] J. Dirutigliano, *Some considerations on the relationship between the right to a reasoned decision and the right to explanation in the proposal of the artificial intelligence act*, The digital constitutionalist the future of constitutionalism, 2023.

[330] The examined example of high-risk AI system is consistent with the characteristics of the public actor presented in chapters 1 and 2, as this work predominantly analyzes the impact of the AI on the duty of the State to evaluate the eligibility of natural persons for essential public assistance benefits and services, including healthcare services, as well as to grant, reduce, revoke, or reclaim such benefits and services.

response services, including by police, firefighters and medical aid, as well as of emergency healthcare patient triage systems; […]'.

In this regard, in addition to the obligations under Art. 9-15, there are specific obligations that have to be fulfilled. The first important measure is the obligation of all public users/deployers to inform affected persons that they are subject to the use of a high-risk system of Annex III. Accordingly, this obligation exists not only where the decision is fully automated, but also where the AI system is used to assist a human in making the decision. It is also specified that this information shall include the intended purpose and the type of decisions the AI system makes. It also obliges to inform the affected person about the right to request an explanation, which will be referred to below[331].

This obligation to inform the concrete affected person is supplemented by the new obligation of only public authorities to register their use of high-risk AI systems in the EU database of high-risk systems envisaged in Art. 71. This is an important transparency obligation for public authorities that will allow control of administrative high-risk systems by public watchdogs and that goes beyond the concrete administrative procedures that must be followed to adopt single-case decisions. Particularly, according to the AI Act only providers are obliged to register AI systems in this database.

Another important measure to protect affected persons is the new right to an explanation envisaged in Art. 86. According to this new provision, deployers (in this case public authorities) that use high-risk AI systems to adopt decisions with legal effects or that adversely affect a natural person must give a clear and meaningful explanation to this affected person when he or she requires it. What has to be explained is the role of the AI system in the decision-making procedure, the main parameters of the decision taken and

[331] Recital 93 of the AI Act.

the related input data. This explanation must only be given on request of the affected person and may be excluded, in justified cases, by Union or Member State law.

In addition, to protect affected parties it has been introduced the right they have to lodge a complaint with the national supervisory authority if they consider that the AI systems relating to them infringe the AI Act. This complaint is without prejudice to any other administrative or judicial remedy that may exist.

Finally, the AI Act provides deployers of high-risk systems to carry out a FRIA prior to their first use. Such an impact assessment must take into account the specific context of use of the AI system and includes the duty to make wide consultations in front of the national supervisory authority and relevant stakeholders, who shall have six weeks to submit comments. Public authorities must publish a summary of this impact assessment when they register the use of the AI system in the aforementioned EU database of high-risk systems. If a data protection impact assessment must also be carried out, it can be included as an addendum to this FRIA.

Lastly, it is important to underline that according to Art. 111, the providers and deployers of high-risk AI systems intended to be used by public authorities shall take the necessary steps to comply with the requirements and obligations of this Regulation by 2 August 2030.

It follows from what has been seen that the AI Act will have a major impact on the public actor. Particularly, it will prohibit some AI systems that many public authorities would want to use. It will impose numerous substantive and procedural obligations on them when developing or using high-risk systems, including obligations to conduct a prior impact assessment with extensive consultation before using the system, to register the use of the system in a European database, and to inform and provide a detailed explanation to natural persons affected by decisions based on such systems[332].

[332] O. Mir Puigpelat, *The impact of the AI Act on public authorities and on administrative procedures,* Rivista Interdisciplinare sul Diritto delle Amministrazioni Pubbliche, 2023.

However, not all the open issues about the protection of fundamental rights within the AI-driven public actor are solved. As such, the following paragraph analyzes them in order to track the new challenges that the interpreters will face during the application of the AI Act.

3.3.6 Left open questions in the protection of fundamental rights

Although the AI Act can be seen as the 'landing point' of the AI regulation, being the first comprehensive piece of legislation that is focused on this specific matter, there are still some issues that the final text seems not to solve specifically in relation to the possible impact on fundamental rights.

Particularly, about prohibited AI systems, Art. 5 defines AI practices that are generally prohibited because their use will cause or is 'likely to cause' to a person or any other person 'physical or psychological harm'. However, the definition of 'harm' it is not clear, as it does not entail explicitly AI systems whose use results in violation of fundamental rights or any other harm, including legal and financial harm. In this regard, the legislator should clarify that a cumulation of (more minor) harms related to AI systems should be regarded as 'harm' within this context.

In relation to transparency obligations, it is not clear the relationship with data protection rights. Particularly, Art. 50 of the AI Act establishes that providers must 'ensure that AI systems intended to interact directly with natural persons are designed and developed in such a way that the natural persons concerned are informed that they are interacting with an AI system, unless this is obvious from the point of view of a natural person who is reasonably well-informed, observant and circumspect, taking into account the circumstances and the context of use'. It is unclear what this provision adds to the already existing transparency requirements in the GDPR, including the obligation of providers to inform data subjects and the individual right to be notified and have access to data being

processed. As such, there is a concern that this transparency provision in the AI Act may be read or applied as a limitation of the general GDPR provisions. The AI act should clarify that none of the provisions in the AI Act can be applied to limit the level of protection offered by the fundamental right to private life and data protection as protected in the CFREU and the GDPR.

Moreover, regarding the FRIA for high-risk AI systems, even if it comes with the best intentions, there is the risk that it will likely remain a paper tiger, generating efforts and costs without much effect. First, it remains unclear to what extent this impact assessment will go beyond the general risk assessment according to Art. 9, which also includes mapping and mitigating risks to health, safety and fundamental rights. A second aspect of attention is that such an assessment is only required for high-risk systems. It is not possible to exclude *a priori* that there may be an infringement of fundamental right also in areas not considered high risk by the AI Act, as it is advisable that this assessment should be conducted for all AI systems affecting natural persons regardless of whether they fall within the areas referred to in Annex III of the AI Act.

Furthermore, safeguard of human intervention should only be guaranteed in cases where automated decisions making produces legal effects concerning the natural person or, similarly, significantly affects the natural person, neither does the AI Act foresee the obligation of human intervention with respect to AI systems that are not classified as 'high risk' to the rights or legitimate interests of data subjects. It follows that in all other cases where automated decision-making might not produce significant effects or a noticeable impact on data subjects, this human intervention should not be considered obligatory by law. In this regard, administrative jurisprudence has not provided guidance on whether it is necessary to ensure human intervention with regard to every type of decision-making

process or only with regard to those that are likely to produce significant effects on data subjects, although it has been pronounced on cases that might fall into the second type[333]. Finally, in order to guarantee the right to effective remedy, a specific provision should be introduced to ensure that national courts and tribunals have sufficient powers and tools to provide effective remedies for individuals. This also means that individuals affected by the application of AI-based tools or goods must have the possibility to address possible flaws in the architecture or development of AI technologies. The necessity of informed decision-making has, albeit in a different context, underlined as being able to ascertain the reasons upon which the decision taken in relation to him or her is based, either by reading the decision itself, or by requesting and obtaining notification of those reasons. Therefore, the inclusion of specific provisions should be proposed in order to ensure the right to information on the development and use of AI, access to information on the development, content and criteria used in an AI-based tool, procedure or scoring system.

4. Conclusive remarks

Ultimately, the use of AI systems by the public actor still raises possible concerns from a fundamental rights perspective, as from a constitutional and legislative view it entails a possible detriment to citizens' rights. This outcome emerges from the 3-layered analysis that has been carried out in this chapter, focusing on the constitutional rights and legislative framework of AI, as well as on the answers provided by the administrative jurisprudence.

On this merit, section 2 of the chapter has analyzed the balance between certain (and selected) fundamental rights and the interest of the public actor in using technologies to

[333] C. Fusco, *The Use of Artificial Intelligence in the Decision-Making Processes of the Public Administration: Regulations and Executive Practice - The Case of the Italian Public Administration*, in (eds.) D. Marino - M. A. Monaca, *Innovations and Economic and Social Changes due to Artificial Intelligence: The State of the Art, Studies in Systems, Decision and Control*, Berlin, 2023.

make the public apparatus more efficient. Following this path, the discussion raised certain questions on whether – and if yes how – the current constitutional European framework is well suited to address the challenges posed by the use of AI systems. As shown, some of these difficulties include, due to the inherent functionalities of the AI systems, as well as the possible use the public actor could make of them, the risk of: (i) violation of the privacy right and the data protection one; (ii) discrimination when prediction analytics come into play; (iii) the allocation of the burden of proof when the State refuses to disclose information about a given AI system in relation to the right to have an effective remedy; (iv) impossibility to guarantee the right to transparency and access to file; (v) impossibility to guarantee a reasoned decision[334].

In addition to this analysis, section 3 examined the responses of the European legislator and national jurisprudence. In relation to the solutions provided by the European legislator, it has been found that the GDPR poses several interpretative problems with respect to its applicability to the use of AI systems. Moreover, it provides for a right to explainability that is too limited in relation to the impact of the legislation on citizens' fundamental rights.

In light of this, the analysis then focused on the responses of national jurisprudence (as examples, a number of rulings from Italian and Dutch administrative jurisprudence have been chosen, as well as of the CJEU). In the absence of specific legislative provisions, judges have so far played a central role in the definition of legal coordinates with respect to the topic of public actor's algorithmic decisions. However, as shown, the path toward building a 'technological due process' within the paradigm of the AI-driven public actor still appears to be on a preliminary stage.

Finally, the analysis focused on the responses given by the legislator in the AI Act, in which not all the open issues about the protection of fundamental rights within the AI-driven

[334] A. Rachovitsa - N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.

public actor are solved even if the Regulation tries to adopt a 'hybrid' legislative approach according to which the technology has to be fundamental right oriented since its design. Accordingly, following this path – making the current provisions more oriented to this approach – is actually the only possible way to ensure that the value (*i.e.*, the meaning) of protecting the fundamental rights of the citizens as their interests become an integral part of the development of AI[335]. Ultimately, the purpose of this chapter is to open a door for AI, without leaving citizens unprotected. This is based on the conviction that technological innovations such as AI cannot be answered unquestioningly with existing legal instruments, but that they regularly demand innovative solutions[336].

[335] A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33.
[336] N. Marsch, A*rtificial Intelligence and the Fundamental Right to Data Protection: Opening the Door for Technological Innovation and Innovative Protection,* in (eds.) T. Wischmeyer – T. Rademacher, *Regulating Artificial Intelligence*, Berlin, 2020.

CHAPTER IV: THE INTERACTION BETWEEN PUBLIC AND PRIVATE ACTORS IN THE 'AI-DRIVEN PUBLIC ACTOR'

Contents: 1. The interaction between public and private actors in the 'AI-driven public actor'; 2. The role of the private actor as enabler of the 'AI-driven public actor'; 2.1 Accountability in AI; 2.2 Processing of public actor's personal data by private parties: the GDPR; 2.3 The externalization of public actor's AI systems within the AI Act; 2.3.1 Qualification of the parties and classification of AI systems; 2.3.1.1 Main interpretative and application issues of the AI Act; 2.3.1.2 EU Model Contractual AI Clauses; 2.3.2 Final remarks on the externalization of public actor's AI systems within the AI Act; 2.4 The allocation of responsibility within the Product Liability Directive; 2.4.1 General aim of the two directives and related concerns; 2.5 Conclusions; 3. Information sharing between private and public actors: a new data sharing model; 3.1 The AI Act: a missed opportunity?; 3.2 Other limits of the actual Business to Government (B2G) data sharing; 3.3 Data Governance Act; 3.3.1 Possible concerns; 3.4 Data Act; 3.5 Open Data Directive; 3.6 Commonalities and elements of consistency within EU data legislation.

1. The interaction between public and private actors in the 'AI-driven public actor'

In order to have a comprehensive view of the AI-driven public actor paradigm, it is essential to understand that the changes that brought to its development are not just strictly linked to the new forms of exercise of public functions, but also to the fundamental role that the private actors are having in this transition.

Two aspects of this strong interconnection between public and private actors can be observed.

First, frequently private actors hold the know-how for the creation and use of sophisticated technologies such as AI. Therefore, the public actor has to rely on the provision of technologies by the private sector which, usually, is not open to sharing proprietaries information about the functioning of the AI systems. Designing algorithmic and machine

learning systems involves decisions about goals, values, risk and certainty, and a choice to place constraints on future agency discretion. If these systems employ adaptive machine learning capabilities, their design choices can have an influence on how public decisions are reached. In this context, when the adoption of those systems is governed by private actors, the risk is that no public participation is permitted as well as no reasoned deliberation, and no factual record. In this sense, a concrete threat exists that public actor responsibility for policymaking is abdicated[337]. Therefore, this aspect shall be investigated specifically in order to understand if the actual European legal framework provides some instruments directed to protect the public actor's interests over the private ones.

Second, due to an inherent close connection between the development of AI and sharing data models exists, the EU is fostering a strategy on data sharing between private and public actors. This in order to enhance the reliability and the generalizability of an AI model (which depend heavily on large amounts of training data that are diverse and representative of the population), as well as to foster transparency within the use of AI between the public and the private actor.

On this matter, the European Commission in the European strategy for data[338] states that:

> 'The value of data lies in its use and re-use. Currently there is not enough data available for innovative re-use, including for the development of artificial intelligence. The issues can be grouped according to who is the data holder and who is the data user, but also depend on the nature of data involved (*i.e.*, personal data, non-personal data, or mixed

[337] D. K. Mulligan – K. A. Bamberger, *Procurement As Policy: Administrative Process for Machine Learning*, Berkeley Technology Law Journal, 34, 2019.

[338] European Commission, *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*, (February 19th, 2020).

data-sets combining the two). Several of the issues concern the availability of data for the public good'[339].

Therefore, in these last years several regulations[340] have been enacted in order to create new mechanisms of data sharing.

In this context, it is interesting to examine how the private sector has taken part in the 'algorithmisation' of the public sector because of the technological dependency that exists between these two actors (section 2). In addition, the regulatory phenomenon concerning data sharing obligations between the public and the private actor is examied (section 3).

2. The role of the private actor as enabler of the 'AI-driven public actor'

As also the public actor understood the potential of AI systems in the public decision processing, a strict dependency between the public and the private actor in the procurement and use of technological solutions can be observed[341]. Indeed, in order to ensure technological development, it is clear that the public sector cannot face the transformation relying solely on internal workforces. As such, it is interesting to note the attention given to public procurement in the Coordinated plan on Artificial Intelligence[342], where it is stated that:

'public procurement is key in public sector AI adoption. It can also help stimulate demand and offer of trustworthy and secure AI technologies in Europe. In this context, the Commission is developing an Adopt AI program to support public procurement of

[339] European Commission, *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*, (February 19th, 2020).
[340] Particularly, the Data Act, the Data governance Act and Open Data Directive have been proposed by the European Union to foster the communication of data and information between the private and public sector.
[341] See Table 1 in Chapter 1, whre the case of eu-LISA is presented.
[342] European Commission, *Coordinated plan on artificial intelligence 2021 review*, *Annexes to the Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Fostering a European approach to Artificial Intelligence*, (April 21st, 2021).

AI systems and help to transform public procurement processes themselves. The program aims to help Europe's public sector to use its strong collective purchasing power to act as a catalyst and stimulate demand for trustworthy AI. The public sector can lead the way in developing, purchasing, and deploying taking in use trustworthy and human-centric AI applications, for example, by utilizing public procurement of innovative solutions or by steering the development of new solutions towards its needs through pre-commercial procurement practices'[343].

The attention brought by the EU is also based on the consideration that governments depend on strong relationships with the private sector, a trend that is likely to intensify as the latter realize the global scope of the AI business opportunity, and governments continue to rely on external technological know-how. As such, public-private collaboration appears consistently as an essential characteristic of public technological development. This dynamic fits within a larger shift towards networked, composite collaboration between the public and the private sector in governance.

Of course, private companies have been entering into legal relationships with public bodies for a long time, both for infrastructure and public service provision. But more recently, this cooperation entails the use of sophisticated systems that are directly involved into the public decision-making processes, as these systems can support the decision makers. Although these agreements do not involve a full transfer of responsibilities from public bodies to the private sector but rather a public-private collaboration with shared responsibilities, they still pose several risks.

In this sense, there is an emerging literature discussing technology procurement in the public sector and the difficulties it presents in terms of safeguarding fundamental rights

[343] *Ibid*.

and other public values[344]. It has been observed that government authorities have no influence on the design of these systems, and thus they lack the ability to steer and align these systems with public values[345]. Because of trade secrecy and confidentiality claims are raised by private vendors[346], public authorities are often unable to exercise any meaningful oversight. In this context, Voorwinden argues 'that private actors do not assume the typical role of a vendor for they can influence policy through the data and models they generate; they can equip and orient emergency services and the police force; they can shape and access civic participation channels; they can provide and develop novel local public services; […] they can monitor and experiment with public space'[347,348]. Moreover, other risks range from discrimination to lack of due process, discontinuance of essential services, and harmful misrepresentations. Additionally, there are challenges in clearly specifying the desired regulatory attributes related to most goals of AI regulations. Most of those attributes are difficult to observe or measure, and the processes leading to their promotion are not easy to establish. The outcomes of those processes are not binary and determining whether a requirement has been met cannot be subject to strict rules, but rather to (yet to be developed) technical standards with an unavoidable degree of undefinition, which may also be susceptible of iterative application in, for example, agile deployment methods, and thus difficult to evaluate at tender stage. The desired attributes can be in conflict between themselves or with the main functional specifications for digital

[344] K. D. Mulligan – K. A. Bamberger, *Procurement As Policy: Administrative Process for Machine Learning*, Berkeley Technology Law Journal, 2019, p. 773; R. Brauneis – E. P. Goodman, *Algorithmic Transparency for the Smart City,* Yale Journal of Law & Technology, 2018, p. 103; A. Voorwinden, *The Privatised City: Technology and Public-Private Partnerships in the Smart City*, Law, Innovation and Technology, 2021, p. 1.
[345] K. D. Mulligan – K. A. Bamberger, *Procurement As Policy: Administrative Process for Machine Learning*, Berkeley Technology Law Journal, 2019, p. 773.
[346] R. Brauneis – E. P. Goodman, *Algorithmic Transparency for the Smart City,* Yale Journal of Law & Technology, 2018, p. 103.
[347] A. Voorwinden, *The Privatised City: Technology and Public-Private Partnerships in the Smart City*, Law, Innovation and Technology, 2021, p. 1.
[348] L. Vandercruysse – A. Christofi – C. Buts – M. Dooms – P. Valcke, *Data Protection in Smart Cities*, European Procurement & Public Private Partnership Law Review, 2022, p. 81 – 93.

technology deployment. There is, for instance, a growing understanding of the incompatibility (or unavoidable trade-off) between requirements for explainability and AI performance, in the sense that non-explainable AI solutions tend to have higher levels of functional performance. Negotiating those trade-offs is complex and subject to nontechnical decisions (*e.g.*, how much more accurate must a solution be to justify a reduction in explainability?). Additionally, other incompatibilities or tensions between goals of digital regulation may be more difficult to identify and balance out, especially if the results of a technological deployment can only be observed and assessed with a significant time lag. This creates a need for the public actor to create solutions through negotiations within the procurement process, or to leave some issues to co-decision at the contract execution phase. While seeking to increase flexibility and to leverage the technical expertise of the tech provider, such approaches also generate significant risks of commercial determination of the content of the regulatory contract[349].

These risks are neither hypothetical nor intangible. Today, AI systems help governments to take different public relevant decisions (*e.g.*, how many hours of care disabled individuals will receive[350]; which employees should be hired, fired, or promoted[351]). Yet as decision-making shifts from human-only to a mixture of human and algorithm, questions of how to allocate liability for the public actor's actions still remains largely unanswered. The majority of solutions to these concerns have focused on technological or regulatory oversight to address bias, fairness, and due process[352]. And the situation becomes more

[349] A. Sanchez-Graells, *Public Procurement of Artificial Intelligence: Recent Developments and Remaining Challenges in EU Law*, Legal Tech Law Journal, 2024.
[350] C. Lecher, *What Happens When an Algorithm Cuts Your Health Care*, The Verge, 2018.
[351] M. Bogden – A. Rieke, *Help Wanted: An Examination of Hiring Algorithms, Equity, and Bias*, Upturn, 2018.
[352] For a survey of these risks and concerns, see generally S. Barocas – A. D. Selbst, *Big Data's Disparate Impact*, California Law Review, 2016, p. 671 (using the lens of antidiscrimination law to explore bias arising from data mining); D. K. Citron – F. Pasquale, *The Scored Society: Due Process for Automated Predictions*, Washington Law Review, 2014, p. 1 (warning that additional procedural safeguards are necessary for automated prediction systems); D. K. Citron, *Technological Due Process*, Washington Law Review, 2008, p. 1249 (proposing a 'technological due process' model to vindicate procedural values in an era of

concerning when vendors are involved in very high-stake decisions like law enforcement or health and benefit systems[353,354].

However, to date, contractual accountability frameworks have failed to address the larger social and structural aspects of the problems[355,356]. As such, it is time to consider new paradigms for accountability[357]. In other words, the creation of a private legal framework outside any representative mechanism can pose threats to democracy due to the marginalization of citizens and their representatives from public decision processing. This situation shows why it is important to focus on the legislative remedies to solve the issues related to the interactions between these actors[358].

In this context, how to best cooperate with private actors remains an interesting topic that requires further investigation from a legal point of view. As such, the legal provisions the private actor has to comply with (or, at least will have to) are analyzed in the following

automation); K. Crawford – J. Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, Boston College Law Review, 2014, p. 93 (arguing that procedural due process provides a framework for the regulation of big data); D. Gray – D. Citron, *The Right to Quantitative Privacy*, Minnesota Law Review, 2013, p. 62 (raising concerns over the use of algorithmic systems to establish probable cause for law enforcement searches or arrests).

[353] A recent case in the USA was a lawsuit claiming Immigration and Customs Enforcements (ICE) created a 'secret no-release policy' and manipulated the risk assessment algorithm to recommend only one decision. Velesaca v. Decker, 458 F. Supp. 3d 224 (S.D.N.Y. 2020) challenged the automatic and indefinite incarceration virtually all of the thousands of people ICE arrested between 2017 and 2020 for alleged immigration offenses. The algorithm used to recommend an arrestee be released or detained until a hearing was changed in 2015 and again in 2017, removing the ability to recommend release, even for arrestees who posed no threat. The detainees were not subject to due process and never had any change at recourse. The settlement in the case in March 2022 secures the right to a fair release assessment for everyone arrested by ICE in New York. Another example is the District Court of the Hague, 6 March 2020, n. 865, presented in Chapter 2 of this thesis, in which the District Court of Hague found that 'under article 8 of the ECHR, the Netherlands did not strike a fair balance between privacy and the benefits of the use of new technologies to prevent and combat fraud because Syri was 'insufficiently clear and verifiable''.

[354] M. Hickok, *Public procurement of artificial intelligence systems: new risks and future proofing*, AI & Society, 2022.

[355] D. Reisman – J. Schultz – K. Crawford – M. Whittaker, *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability*, AI Now Institute, 2018.

[356] Please note that this statement does not consider the possible positive effects the AI Act and the Product Liability Directive should bring in this context. Indeed, in order to provide a complete view on the actual legal framework, the regulatory choices and remedies identified by the legislator in the AI Act and in the Product Liability Directive are presented in paragraphs 2.4 – 2.4.1.

[357] K. Crawford – J. Schultz, *AI Systems as State Actors*, Columbia Law Review, 2019, p. 1941.

[358] O. Pollicino – G. De Gregorio, *Constitutional Law in the Algorithmic Society*, in (eds.) H. Micklitz - O. Pollicino - A. Reichman - A. Simoncini - G. Sartor - G. De Gregorio, *Constitutional Challenges in the Algorithmic Society*, Cambridge, 2021.

sections which delve into the third parties engagement discipline provided for by the GDPR (section 2.2), the AI Act (section 2.3) the Directive (EU) 2024/2853 of the European Parliament and of the Council of 23 October 2024 on liability for defective products and repealing Council Directive 85/374/EEC (Products Liability Directive) (section 2.4)[359].

2.1 Accountability in AI

The interplay between private and public actors in the procurement of AI technologies is first and foremost linked to the concept of accountability, from which any liability stems. Indeed, the analysis of the various regulations proposed in the following paragraphs is aimed precisely at highlighting the relationship between accountability and liability, in order to determine the responsibilities of the parties in the context of the use by the public actor of AI systems provided by private actors.

Accountability is one of the cornerstones of the AI governance. This is, among other reasons, because of the delegation of tasks (*e.g.*, prediction or decision-making) to AI systems. Current AI policy, particularly in the European context, recognizes that if AI is to be used to assist or delegate decision-making, it is necessary to ensure that these systems are fair in their impact on human lives, consistent with values that should not be compromised, capable of acting accordingly, as well as that appropriate accountability processes can be ensured.

Accountability is often largely undefined. This is clear in some of the major European documents on AI, such as the GDPR, the AI Act and the Product Liability Directive. In the GDPR, accountability is defined both as a principle that ensures compliance with the key requirements for a trustworthy AI and as a set of practices and measures, *e.g.*, audit, risk management, and redress for adverse impact (Art. 5 and 24). Here, accountability works

[359] L. Torchia, *Lo stato digitale*, Il Mulino, 2023.

as a meta-principle directed at data controllers so that they demonstrate, by virtue of their information background, compliance with GDPR requirements in the processing of personal data and as a remedy mechanism for failure to comply with them: 'The controller shall be responsible for, and be able to demonstrate compliance with fairness, transparency, purpose limitation, data minimization, storage limitation, accuracy, confidentiality'[360]. Moreover, the AI Act also contains an undefined concept of accountability, aligned with the risk-based regulatory approach: providers and deployers of AI are accountable for different reasons and in different ways depending on the risk level of the respective AI systems[361]. Finally, the Product Liability Directive does not provide any definition of accountability.

Unfortunately, an imprecise definition of accountability is problematic, not least because it risks undermining the public debate and policy-making. This is challenging especially when political and legislative agreements have not yet been formed, including the accountability for the provision of many AI services[362].

2.2 Processing of public actor's personal data by private parties: the GDPR

As previously mentioned, in the context of the use of AI by the public actor the trend toward outsourcing of services concerning AI technologies to private entities cannot be ignored.

In this regard, it is likely that these procurements are linked to the use of personal data collected by the public actor to train the private proprietaries algorithms. Particularly, the public actor, as the data controller of this personal data, should ensure that it is lawfully processed under the GDPR. Therefore, it is interesting to analyze the GDPR provisions

[360] Art. 5(1) and (2) of the GDPR.
[361] C. Novelli – M. Taddeo – L. Floridi, *Accountability in artificial intelligence: what it is and how it works*, AI & SOCIETY, 2023, p. 1-12.
[362] *Ibid*.

in order to understand which safeguards are put in place in the case the public actor is qualified as the data controller, whereas the supplier of processing services is qualified as data processor[363].

In this context, a better understanding of the concept of data controller and data processor plays an essential role in the application of the GDPR, as the specific roles of the parties determine their responsibilities for the processing of data subjects' personal data. Under the GDPR, a controller is defined as a 'the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data'[364]. A processor, in turn, is 'a natural or legal person, public authority, agency or other body which processes personal data on behalf of the controller'[365].

In the use of AI, data sets can become overwhelmingly big and scattered. Moreover, the use of these technologies and the linked processing of personal data can result in difficulty for public officials, who lacking technical expertise must turn to third parties. As a result, public data controllers often outsource the processing to several data processors specialized in processing masses of data. For the same reason, it is common that the initial processors further delegate parts of the processing to sub-contractors (sub-processors).

Under Art. 28(1) of the GDPR the controller must 'use only processors providing sufficient guarantees to implement appropriate technical and organizational measures in such a manner that processing will meet the requirements of this Regulation and ensure the protection of the rights of the data subject'. This means that the currently applicable

[363] Please note that only the data controller (public actor) – data processor (private actor) relationship is considered in this case, as it results useful to illustrate the application of the GDPR. However, it cannot be excluded that, based on practical circumstances, the existence of joint-controllership relationship, for example, does not emerge.
[364] Art. 4, no. 7 of the GDPR.
[365] Art. 4, no. 8 of the GDPR.

legislation requires ‘pre-contractual checks’ on the processors. Moreover, Art. 28(3) provides for a series of different obligations as the processing by a processor shall be governed by a contract or other legal act under Union or Member State law, that is binding on the processor and that sets out the subject-matter and duration of the processing, the nature and purpose of the processing, the type of personal data and categories of data subjects and the obligations and rights of the controller[366]. These obligations include, *inter alia*, the duty of the processor to:

a. processes the personal data only on documented instructions from the controller, including with regard to transfers of personal data to a third country or to an international organization. Thus, the public actor must give specific instructions on the processing of personal data indicating, for example, the specific purposes for which they can be processed, as well as the means the processor can use. However, a tension exists between the use of AI and the purpose limitation requirement. These technologies enable the useful reuse of personal data for new purposes that are different from those for which the data have been originally collected[367]. Moreover, in view of the information asymmetry that exists between the public and private actors, it could be the case that the data processor does not comply with the instructions received, as well as that the public actor does not have any effective means of control;
b. to keep detailed documentation about the processing chain and requires processors to obtain authorization from the data controller when engaging sub-processors, thus adding a further control mechanism to protect data subjects’

[366] J. Lindqvist, *New Challenges to Personal Data Processing Agreements: Is the GDPR Fit to Deal with Contract, Accountability and Liability in a World of the Internet of Things*, International Journal of Law and Information Technology, 2018, p. 45.
[367] Panel for the Future of Science and Technology EPRS, *The impact of the General Data Protection Regulation (GDPR) on Artificial intelligence*, European Parliamentary Research Service Scientific Foresight Unit (STOA), (June 2020).

rights. The reasons why transparency in sub-processor contracting relationships is important are twofold. First, it is important from a contractual point of view for the data subjects' legal security to be able to identify which party is liable to him or her in the event of damage or loss. Secondly, it is important for the data controller to be aware of all the engaged sub-processors to be able to ensure compliance with the legal requirements provided by the GDPR[368];

c. at the choice of the controller, deletes or returns all the personal data to the controller after the end of the provision of services relating to processing, and deletes existing copies unless Union or Member State law requires storage of the personal data.

In this context, it should be noted that the programmers of modern machine learning systems create data sets to be used as training data. On the basis of this data set the machine is requested to run the algorithm on the training data and to achieve a given goal by finding common patterns and producing a model that can be further deployed to achieve the ultimate goal and outcome. Hence, in this complex pattern, some questions emerge. When the data processor receives a request for erasure from the data subject, which kind of data is to be deleted? From which datasets? And, more importantly, how to obtain the erasure? Notably, in the case of different datasets, it is natural that personal data may be involved both in the training set and in the analysis set.

It is to be said now that there are no crystal-clear answers to these questions. Some technical remedies are proposed, such as anonymizing data, functional encryption, selective amnesia, and model breaking. However, none of them seems to tackle

[368] J. Lindqvist, *New Challenges to Personal Data Processing Agreements: Is the GDPR Fit to Deal with Contract, Accountability and Liability in a World of the Internet of Things*, International Journal of Law and Information Technology, 2018, p. 45.

the core of the problem: clarify the extent of the right to erasure from the machine learning perspective. A huge issue is to be found in the current lack of legal certainty as to how AI can be designed to comply with the regulation, given the specific features of data protection rights. Moreover, this tension between GDPR and machine learning happens since the latter are designed to render data (unilateral) modification difficult. This matter is hard to reconcile with the GDPR's requirement that personal data be erased when specific circumstances apply. Hence, three main relevant conceptual uncertainties threaten data subjects' rights and processors' obligations.

First and foremost, many uncertainties rely on the term 'erasure'. Deleting data from machine learning data sets is burdensome since it implies retraining the entire model. It does not address the underlying problem of making sensitive data disappear or become untraceable. Secondly, it is challenging to demonstrate that the retrained model is fully 'corrected'. Namely, it has been cleaned up from the wrongly obtained data, and the biased ones are not reproduced. Technical factors and governance design thus burden the difficulty of complying with Art. 17 of the GDPR. Indeed, even if there would be a means of ensuring compliance from a technical perspective, reaching out to all the datasets may be organizationally tricky. Thirdly, because of a certain degree of unpredictability and autonomy, it is frequently challenging to find the liable party in the case of damage caused by AI applications. In particular, the more challenging situations are those in which the outcome of the processing carried out by the AI is not fully controllable *a priori*. Moreover, according to the principle of accountability is the processor's duty to 'take into account state of the art, the costs of implementation and the nature, scope, context, and purposes of the processing, as well as the risk of varying likelihood

and severity for the rights and freedoms of natural persons. Therefore, the controller, and the processor, shall implement appropriate technical and organizational measures to ensure a level of security which can be considered appropriate to the risk'[369].

Hence, the legislator delegates to the data processor – and ultimately to the data controller – the burden of identifying how to fulfil the requirements dictated by the rule, dropping them into the concrete case, and taking responsibility not only for implementation but also for evaluating the risks. Those aspects emerge when the processing is not linear and involve data controllers and several sub-controllers since, often, their contracts establish the execution of some data subject's rights, including the right to erasure. Therefore, the logic of accountability is challenged not only by the crowded level of responsibilities arising from the regulation but also in the case of assigning responsibilities to the presence of automated decision-making[370].

These issues of compatibility between the use of AI systems and GDPR are made even more difficult to solve by the circumstances that in theory controllers determine how and why processing takes place and are largely accountable for compliance with the legislative provisions. Processors act under their instruction and on their behalf. However, the increasingly complex, networked, and dynamic nature of contemporary processing environments challenges this understanding of roles and responsibilities and, thus, data protection law's ability to effectively protect data subject. The dominant-subordinate understanding of controller-processor relationships, still depicted in GDPR, does not reflect the actual power dynamics in many contemporary environments. Though

[369] Art. 32 of the GDPR.
[370] F. Paolucci, *The costs of training AI and the impact on fundamental rights*, The Digital Constitutionalist the Future of Constitutionalism, 2023.

customers are sometimes afforded some choice, turn-key AI services are offered generically as determined by the provider's standard form contracts. Large companies with technical expertise and market power will of course use AI services, perhaps on a consultancy basis involving close cooperation with the provider, though still typically through the provider's specified contractual arrangements (effectively positioning providers to influence activities of the public sector)[371].

In this context, neither the accountability regime the data processor can be subject to seems to solve the imbalance of powers between the public and the private actor in the processing of personal data for AI purposes. Indeed, Art. 82 of the GDPR brings changes to the liability distribution stating that 'any person who has suffered material or non-material damage as a result of an infringement of this Regulation shall have the right to receive compensation from the controller or processor for the damage suffered'. According to Art. 82(2) of the GDPR, a controller is liable for damage caused by processing which infringes the regulation, whilst a processor 'shall be liable for the damage caused by processing only where it has not complied with obligations of this Regulation specifically directed to processors or where it has acted outside or contrary to lawful instructions of the controller'[372]. The underlying aim is to assure full and effective compensation for damages on data subjects. Indeed, the first and foremost role of the controller is to allocate responsibility and to determine who is responsible for compliance rules, and how data subjects can exercise their rights in practice[373]. However, as noted above, the supervision that the public actor, as data controller, should exercise is defeated by the asymmetries of

[371] J. Cobbe – J. Singh, *Artificial intelligence as a service: Legal responsibilities, liabilities, and policy challenges*, Computer Law & Security Review: The International Journal of Technology Law and Practice, 2021.
[372] Art. 82(2) GDPR.
[373] J. Lindqvist, *New Challenges to Personal Data Processing Agreements: Is the GDPR Fit to Deal with Contract, Accountability and Liability in a World of the Internet of Things*, International Journal of Law and Information Technology, 2018, p. 45.

information as well as by technical deficiencies, which result also in an imbalance of powers between the two contractual parties.

Therefore, in this context the risk that the private actor's processing activities could be beyond the control of the public actor is even more compelling.

2.3 The externalization of public actor's AI systems within the AI Act

Considering the gaps left open by the GDPR in its application in AI-driven framework, it is interesting to examine the discipline established in the AI Act in the case of public actor procurement of AI systems.

Particularly, the following paragraphs analyze which role the public actor could have in the procurement of AI under the AI Act, as well as the classification of the AI system (section 2.3.1) and some final remarks on the externalization of public actor's AI systems within the AI Act (section 2.3.2).

2.3.1 Qualification of the parties and classification of AI systems

In this context, the first interpretive challenge is to correctly qualify the parties among the AI procurement chain. This is because the AI Act establishes different obligations due to the role assigned to the parties. Particularly, they can qualify as:

- providers, means a natural or legal person, public authority, agency or other body that develops an AI system or a general-purpose AI model or that has an AI system, or a general-purpose AI model developed and places it on the market or puts the AI system into service under its own name or trademark, whether for payment or free of charge. The rules established by the AI Act apply to providers placing on the market or putting into service AI systems in a non-discriminatory manner, namely irrespective of whether the said providers are physically present or established

within the EU or in a third country, and to users of AI systems who are physically present or established within the EU. They also apply to providers of AI systems who are physically present or established in a third country, to the extent the output produced by those systems is used in the EU[374];

- deployers, means a natural or legal person, public authority, agency or other body using an AI system under its authority except where the AI system is used in the course of a personal non-professional activity;
- authorized representative means a natural or legal person located or established in the Union who has received and accepted a written mandate from a provider of an AI system or a general-purpose AI model to, respectively, perform and carry out on its behalf the obligations and procedures established by this Regulation;
- importers, means a natural or legal person located or established in the Union that places on the market an AI system that bears the name or trademark of a natural or legal person established in a third country. The AI Act applies to importers of AI systems, to product manufacturers placing on the market or putting into service an AI system together with their product and under their own name or trademark[375];
- distributors means a natural or legal person in the supply chain, other than the provider or the importer, that makes an AI system available on the Union market. The AI Act applies to distributors of AI systems, to product manufacturers placing on the market or putting into service an AI system together with their product and under their own name or trademark[376];
- operators means a provider, product manufacturer, deployer, authorized

[374] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.
[375] *Ibid*.
[376] *Ibid*.

representative, importer or distributor.

Moreover, in order to understand the possible obligation to be applied also the classification of the AI system should be assessed[377]. Specifically, the AI Act identifies:

- prohibited AI practices: the AI Act lays down a ban on a limited set of uses of AI. In particular, Chapter 2 of the AI Act covers all those AI systems whose use is considered unacceptable as contradicting EU values, especially by violating fundamental rights. As the Commission put it, those specific use of AI systems could be 'harmful and abusive and should be prohibited because they contradict Union values of respect for human dignity, freedom, equality, democracy and the rule of law and fundamental rights enshrined in the Charter, including the right to non-discrimination, to data protection and to privacy and the rights of the child'[378]. The prohibition covers placement on the market, putting into service or use of certain AI systems intended to distort human behavior, whereby physical or psychological harm is likely to occur. These intrusive practices include AI used for adverse behavioral influencing, social scoring, and large-scale surveillance. In particular, Art. 5 of the AI Act explicitly bans harmful AI practices that are considered to be a threat to people's safety and rights, because of the 'unacceptable risk' they create. Accordingly, for example, it would be prohibited to place on the market, put into services or use in the EU: (1) AI systems that deploy harmful manipulative 'subliminal techniques'; (2) AI systems that exploit specific vulnerable groups (physical or mental disability); (3) AI systems used by public authorities, or on their behalf, for social scoring purposes; (4) 'Real-time' (remote)

[377] For a more detailed analysis please see chapter III of this work.
[378] Recital 28 of the AI Act.

biometric identification systems in publicly accessible spaces for law enforcement purposes, except in a limited number of cases[379];

- high risk uses: Chapter III of the AI Act identifies a separate tier of high-risk AI systems and contains specific rules for those AI systems that cause a high risk to the health and safety or fundamental rights of natural persons. In line with a risk-based approach, those high-risk AI systems should be permitted on the European market on a restricted basis, with specific controls in place to support safe use and subject to compliance with certain mandatory requirements and an *ex-ante* conformity assessment[380]. Particularly, AI system shall be considered to be high-risk where: a) the AI system is intended to be used as a safety component of a product, or the AI system is itself a product, covered by the Union harmonization legislation listed in Annex I; b) the product whose safety component pursuant to point (a) is the AI system, or the AI system itself as a product, is required to undergo a third-party conformity assessment, with a view to the placing on the market or the putting into service of that product pursuant to the Union harmonization legislation listed in Annex I. In addition, AI systems referred to in Annex III of the AI Act shall be considered to be high-risk (*e.g.*, AI systems used for biometrics, critical infrastructure, education and vocational training, employment, workers' management and access to self-employment, access to and enjoyment of essential private services and essential public services and benefits, law enforcement, migration, asylum and border control management, administration of justice and democratic processes);

[379] N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.
[380] *Ibid*.

- AI systems intended to interact directly with natural persons: these systems shall be designed and developed in such a way that the natural persons concerned are informed that they are interacting with an AI system, unless this is obvious from the point of view of a natural person who is reasonably well-informed, observant and circumspect, taking into account the circumstances and the context of use;
- general-purpose AI model: an AI model, including where such an AI model is trained with a large amount of data using self-supervision at scale, that displays significant generality and is capable of competently performing a wide range of distinct tasks regardless of the way the model is placed on the market and that can be integrated into a variety of downstream systems or applications, except AI models that are used for research, development or prototyping activities before they are placed on the market.

Depending on the role and the risk level of the AI system, the AI Act imposes different obligations and responsibilities on these actors. For example, providers of high-risk AI systems have to comply with the requirements and obligations for such systems, such as data quality, human oversight, transparency, accuracy, robustness and security. They also have to register their systems in the EU database, conduct a conformity assessment, draw up technical documentation and ensure post-market monitoring. While deployers of high-risk AI systems have *e.g.*, control and risk-management, notification and transparency obligations. Providers of transparency-risk AI systems have to ensure transparency for certain systems, such as chatbots, deepfakes, or emotion recognition. Providers of general-purpose AI models, which are AI systems that can be used for multiple purposes, have to comply with specific obligations, such as providing information and documentation, respecting copyright law and making publicly available a summary of the content used for training.

2.3.1.1 Main interpretative and application issues of the AI Act

One key area requiring clarification in the AI Act pertains to the qualification of the actors subject to its obligations. Concretely, the Regulation mainly differentiates its obligations between deployers and providers. Particularly, the distribution of obligations follows a 'shared responsibility' model: providers and deployers bear distinct responsibilities to prevent undue burden on a single party, at least in principle. A provider is broadly defined as any entity developing an AI system and introducing it to the market or using it in service. The Regulation imposes stringent obligations on providers, particularly for high-risk AI systems, emphasizing compliance with European legal requirements and accountability principles. Conversely, a deployer refers to an actor utilizing an AI system under its authority, excluding non-professional personal activities. Deployers generally need to ensure CE conformity. Also, deployers assume greater responsibility when they are altering the foundational model of an AI system.

Accordingly, determining who qualifies as a provider or deployer poses a legal challenge since the differentiation depends on the interpretation of the legal standard, defined as 'substantial modification' of the system, as determined in Art. 3(1) n. 23. The aim of the AI Act is to bear deployers' responsibility on AI only if they alter the foundation models, but the Regulation does not provide clearly which modifications suffice the threshold[381]. Hence, determining the role of the public actor within the procurement process can result sufficiently evident specifically when it carries out technical intervention on the AI model. These interpretive challenges create problems at the application level, especially in the context of qualifying the relationship between the public actor and the private actor providing AI systems.

[381] F. Paolucci, *Shortcomings of the AI Act: Evaluating the New Standards to Ensure the Effective Protection of Fundamental Rights*, VerfBlog, 2024.

In addition, even if it is possible to correctly qualify the parties, it should not be underestimated that most of the public actor's uses of AI systems may result outside of the application scope of the AI Act, not being classified in none of the categories above presented (section 2.3.1). Indeed, under the AI Act, public sector AI use will trigger complex and fundamental rights-oriented obligations only where such use is classified as 'high-risk' and if the use cannot be exempted.

This creates two threshold issues that can significantly limit the scope of application of the AI Act in relation to public sector digitalization. First, there is the threshold issue of whether a given digital technology solution is to be classed as AI system under art. 3(1) n. 1 of the AI Act[382]. Second, there is the threshold issue of establishing what constitutes a non-exempt high-risk use of AI by the public sector. There are two elements of relevance to this assessment. A first relevant issue is that Art. 6(2) of the AI Act refers to Annex III for a determination of which uses are high-risk. The list in Annex III includes several public sector's AI uses, but the list is by no means exhaustive. Setting aside for now the relatively more clear-cut coverage of biometric identification systems, applications to critical infrastructure, education, law enforcement, or migration, asylum, and border control management (which analysis exceeds the possibilities of this paper), the bulk of public sector AI uses covered in the Annex III presents interpretative challenges. Category 5 refers to access to and enjoyment to 'essential public services and benefits'. These are undefined in the Act, but Recital 58 offers some guidance in indicating that:

> 'another area in which the use of AI systems deserves special consideration is the access to and enjoyment of certain essential private and public services and benefits necessary for people to fully participate in society or to improve one's standard of living.

[382] Under Art. 3(1) n. 1 of the Act, 'AI system' means a machine-based system that is designed to operate with varying levels of autonomy and that may exhibit adaptiveness after deployment, and that, for explicit or implicit objectives, infers, from the input it receives, how to generate outputs such as predictions, content, recommendations, or decisions that can influence physical or virtual environments.

In particular, natural persons applying for or receiving essential public assistance benefits and services from public authorities namely healthcare services, social security benefits, social services providing protection in cases such as maternity, illness, industrial accidents, dependency or old age and loss of employment and social and housing assistance, are typically dependent on those benefits and services and in a vulnerable position in relation to the responsible authorities. If AI systems are used for determining whether such benefits and services should be granted, denied, reduced, revoked or reclaimed by authorities, including whether beneficiaries are legitimately entitled to such benefits or services, those systems may have a significant impact on persons' livelihood and may infringe their fundamental rights, such as the right to social protection, non-discrimination, human dignity or an effective remedy and should therefore be classified as high-risk. Nonetheless, this Regulation should not hamper the development and use of innovative approaches in public administration, which would stand to benefit from a wider use of compliant and safe AI systems, provided that those systems do not entail a high risk to legal and natural persons. In addition, AI systems used to evaluate the credit score, or creditworthiness of natural persons should be classified as high-risk AI systems, since they determine those persons' access to financial resources or essential services such as housing, electricity, and telecommunication services. AI systems used for those purposes may lead to discrimination between persons or groups and may perpetuate historical patterns of discrimination, such as that based on racial or ethnic origins, gender, disabilities, age or sexual orientation, or may create new forms of discriminatory impacts. However, AI systems provided by Union law for the purpose of detecting fraud in the offering of financial services and for prudential purposes to calculate credit institutions' and insurance undertakings' capital requirements should not be considered to be high-risk

under this Regulation. Moreover, AI systems intended to be used for risk assessment and pricing in relation to natural persons for health and life insurance can also have a significant impact on persons' livelihood and if not duly designed, developed and used, can infringe their fundamental rights and can lead to serious consequences for people's life and health, including financial exclusion and discrimination. Finally, AI systems used to evaluate and classify emergency calls by natural persons or to dispatch or establish priority in the dispatching of emergency first response services, including by police, firefighters and medical aid, as well as of emergency healthcare patient triage systems, should also be classified as high-risk since they make decisions in very critical situations for the life and health of persons and their property'.

Category 5 in Annex III then lists specific essential public services covered, and does so in ways that do not entirely match the aforementioned recital:

a. AI systems intended to be used by public authorities or on behalf of public authorities to evaluate the eligibility of natural persons for essential public assistance benefits and services, including healthcare services, as well as to grant, reduce, revoke, or reclaim such benefits and services;
b. AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score, with the exception of AI systems used for the purpose of detecting financial fraud;
c. AI systems intended to evaluate and classify emergency calls by natural persons or to be used to dispatch, or to establish priority in the dispatching of emergency first response services, including by police, firefighters and medical aid; as well as of emergency healthcare patient triage systems.

Quite a few uses of AI in the general functioning of the public sector are thus covered. Remarkably, high-risk classification is only triggered where the system directly affects a

natural person. This sets aside vast areas of administrative activity in economic law with potentially significant effects on individuals (*e.g.*, in relation to micro and small enterprises), as well as AI uses in strictly internal processes. Moreover, Art. 6(3) of the AI Act foresees that:

> 'AI systems shall not be considered as high risk if they do not pose a significant risk of harm, to the health, safety or fundamental rights of natural persons, including by not materially influencing the outcome of decision making. This shall be the case if one or more of the following criteria are fulfilled: (a) the AI system is intended to perform a narrow procedural task; (b) the AI system is intended to improve the result of a previously completed human activity; (c) the AI system is intended to detect decision-making patterns or deviations from prior decision making patterns and is not meant to replace or influence the previously completed human assessment, without proper human review; or (d) the AI system is intended to perform a preparatory task to an assessment relevant for the purpose of the use cases listed in Annex III. [But] an AI system shall always be considered high- risk if the AI system performs profiling of natural persons'.

Given the complexity and relative fuzziness of some of these criteria, the Commission is tasked with providing additional guidance within 18 months from the entry into force of the AI Act.

While the extent to which this will potentially create problems in practice depends on the content of such future guidance, it seems clear that many cases of AI adoption by the public sector could be susceptible of exemption and, more importantly, that there is a risk that public buyers and tech providers carry out the (self)assessment of whether a specific use is high-risk in ways that do not necessarily err on the side of caution. This is a significant flaw of the approach taken in the AI Act as there will be large numbers of AI

deployments classified as non-high risk, either due to the narrowness of the category 5 of Annex III, or as a result of an exemption.

The majority, or at least a very significant number of AI procurements will thus not be covered by the regulatory requirements of the AI Act. This will perpetuate the current scenario of under-regulation. Moreover, there are open questions as to whether Member States can take measures to remedy this situation within their jurisdiction[383]. This issue of the narrow application of the requirements of the AI Act for high-risk uses has an immediate knock-on effect on the likely effectiveness of the EU's model contractual clauses for the procurement of AI[384].

2.3.1.2 EU AI Clauses

In October 2023, the European Commission published the final version of the EU model contractual AI clauses to pilot in procurements of AI[385], which have been developed with the aim to establish responsibilities for trustworthy, transparent, and accountable development of AI technologies by public organizations.

The model AI clauses seek to allow public buyers to ensure compliance with the AI Act by cascading the relevant obligations and requirements down to tech providers (largely on a back-to-back basis). By the same regulatory logic, this technique will be a conveyor belt for the shortcomings of the AI Act, which will be embedded in public contracts using the clauses. It is thus important to understand the shortcomings inherent to this approach and to the model AI clauses, before assuming that their use will actually ensure the trustworthy,

[383] A. Sanchez-Graells, *Public Procurement of Artificial Intelligence: Recent Developments and Remaining Challenges in EU Law,* Legal Tech Journal, 2024.
[384] *Ibid*.
[385] European Commission Public Buyers Community on the Procurement of AI, *New version of Procurement Clauses of AI available: supporting responsible use of AI in Public Authorities,* (October 5th, 2023).

transparent, and accountable development and deployment of AI technologies. Much more is needed than mere reliance on the model AI clauses[386].

As established in the previous section 2.3.1.1, the AI Act will not be applicable to all types of AI use. Remarkably, most requirements are limited to 'high-risk' AI uses as defined in its Art. 6. This immediately translates into the generation of two sets of model AI clauses: one for 'high-risk' AI procurement, which embeds the requirements expected to arise from the AI Act once finalized (the 'high-risk' model)[387], and another 'light version' for non-high-risk AI procurement[388], which would support the voluntary extension of some of those requirements to the procurement of AI for other uses, or even to the use of other types of algorithmic solutions not meeting the regulatory definition of AI.

A first observation is that the complexities surrounding the definition of 'high-risk' AI use by the public sector in the AI Act immediately carries over to the model AI clauses and to the choice of 'demanding' vs light version. As the final version of the AI Act effectively embed a self-assessment of what uses are bound to be high-risk by reference to the relevant exemptions, there are clear risks of gaming the self-assessment to be carried out by tech providers to avoid compliance with the heightened obligations under the AI Act (and it is unclear that the system of oversight and potential fines foreseen in the AI Act will suffice to prevent this). This will directly translate into a risk of gaming (or strategic opportunism) in the choice between 'demanding' vs light version of the model AI clauses by public buyers as well. As mentioned above, it seems that most procurement of AI will be subject to the light version of the model AI clauses, where contracting authorities will need to decide which clauses to use and which standards to refer to. Importantly, the light

[386] Similar issues arise in relation to other sets of proposed clauses. For discussion, see A. Sanchez-Graells, *More Model Contractual AI Clauses – Some Comments on the SCL AI Clauses*, How to crack a nut, (October 18th, 2023).
[387] European Commission Public Buyers Community on the Procurement of AI, *EU model contractual AI clauses to pilot in procurements of AI*, (September 29th, 2023).
[388] *Ibid*.

version does not include default options in relation to quality management, conformity assessments, corrective actions, inscription in an AI register, or compliance and audit (some of which are also optional under the 'demanding' model). This means that, unless public buyers are familiar with both sets of model AI clauses, taking the light version as a starting point already generates a risk of under-inclusiveness and under-regulation that does not align well with the general shortcomings in the regulation of public sector AI use by contract discussed above. Indeed, the model AI clauses come with some additional 'caveat emptor' warnings. They contain provisions specific to AI systems and on matters covered by the proposed AI Act, thus excluding other obligations or requirements that may arise under relevant applicable legislation such as the GDPR. Furthermore, these EU model contractual AI clauses do not comprise a full contractual arrangement. They need to be customized to each specific contractual context. For example, EU model contractual AI clauses do not contain any conditions concerning intellectual property, acceptance, payment, delivery times, applicable law or liability. The EU model contractual AI clauses are drafted in such a way that they can be attached as a schedule to an agreement in which such matters have already been laid down[389]. This is an important warning, as the sole remit of the model AI clauses links back to the AI Act and, in the case of the light version, only partially.

Beyond that, the most significant shortcoming of the model AI clauses is that, by design, they do not include any substantive or material constraints or requirements on the development and use of AI. All substantive obligations are meant to be incorporated by reference to the future (harmonized) standards to be developed under the AI Act, other sets of standards or, more generally, the state of-the-art. Plainly, there is no definition or requirement in the model AI clauses that establishes the meaning of *e.g.*, 'trustworthiness'

[389] European Commission Public Buyers Community on the Procurement of AI, *New version of Procurement Clauses of AI available: supporting responsible use of AI in Public Authorities,* (October 5th, 2023).

and there is thus no baseline safety net ensuring it. Similarly, most requirements are offloaded to (yet to emerge) standards or the technical and organizational measures devised by the parties. For example:

- obligations on record-keeping (Art. 5 high-risk model) refer to capabilities conforming 'to state of the art and, if available, recognized standards or common specifications';
- measures to ensure transparency (Art. 6 high-risk model) are highly qualified: 'The Supplier ensures that the AI System has been and shall be designed and developed in such a way that the operation of the AI System is sufficiently transparent to enable the Public Organization to reasonably understand the system's functioning'. Moreover, the detail of the technical and organizational measures that need to be implemented to reach those (qualified) goals is left entirely undefined in the relevant Annex (E), thus leaving the option open for referral to emerging transparency standards;
- measures on human oversight (Art. 7 high-risk model) are also highly qualified: 'The Supplier ensures that the AI System has been and shall be designed and developed in such a way, including with appropriate human-machine interface tools, that it can be effectively overseen by natural persons as proportionate to the risks associated with the system'. Although there is some useful description of what 'human oversight' should mean as a minimum (Art. 7(2)), the details of the technical and organizational measures that need to be implemented to reach those (qualified) goals is also left entirely undefined in the relevant Annex (F), thus leaving the option open for referral to emerging 'human on the loop' standards;
- measures on accuracy, robustness, and cybersecurity (Art. 8 high-risk model) follow the same pattern. Annexes (G) and (H) on levels of accuracy and on

measures to ensure an appropriate level of robustness, safety and cybersecurity are also blank. While there can be mandatory obligations stemming from other sources of EU law (*e.g.*, the NIS 2 Directive), only partial aspects of cybersecurity will be covered, and not in all cases;

- measures on the 'explainability' of the AI (Art. 13 high-risk model) fall short of imposing an absolute requirement of intelligibility of the AI outputs, as the focus is on a technical explanation, rather than a contextual or intuitive explanation.

All in all, the model AI clauses are primarily an empty regulatory shell. Operationalizing them will require reliance on future (harmonized) standards (*e.g.*, on transparency, human oversight, accuracy, explainability, etc. or, most likely (at least until such standards are in place) significant additional concretization by the public buyer seeking to rely on the model AI clauses.

For the reasons identified above, this is likely to generate regulatory tunnelling and to give the upper hand to AI providers in making sure they can comfortably live with requirements in any specific contract. The regulatory tunnelling stems from the fact that all meaningful requirements and constraints are offloaded to the (harmonized) standards to be developed. And it is no secret that the governance of the standardization process falls well short of ensuring that the resulting standards will embed high levels of protection of the desired regulatory goals, some of which are very hard to define in ways that can be translated into procurement or contractual requirements anyway. Moreover, public buyers with limited capabilities will struggle to use the model AI clauses in ways that meaningfully establish responsibilities for trustworthy, transparent, and accountable development and deployment of AI technologies, other than in relation to those standards.

The content of the all too relevant schedules in the model AI clauses will either simply refer to emerging standards or where there is no standard or the standard is for whatever

reason considered inadequate, be left for negotiation with tech providers, or be part of the evaluation (*e.g.*, tenderers will be required to detail how they propose to regulate accuracy). Whichever way this goes, this puts the public buyer in a position of rule-taker. In this regard, only very few, well-resourced, highly skilled public buyers (if any) would be able to meaningfully flesh out a comprehensive set of requirements in the relevant annexes to give the model AI clauses sufficient bite. And they would not benefit much from the model of AI clauses as it is unlikely that in their sophistication they would not have already come up with similar solutions.

Therefore, at best, the contribution of the model AI clauses is rather marginal, and at worst, it comes with a significant risk of regulatory complacency. This leads to the conclusion that the bulk of the practical regulatory challenges arising from the procurement and deployment of AI by the public sector remain partially unaddressed[390].

2.3.2 Final remarks on the externalization of public actor's AI systems within the AI Act

The analysis above has shown that the EU's regulatory strategy in relation to public sector AI procurement and deployment presents many gaps.

First, significant amounts of AI procurements and use cases will remain unregulated even under the AI Act. Second, even those 'high-risk' use cases covered by the Act will be subjected to mechanisms of regulation by contract that can hardly prove effective given the structural limitations in using public procurement as a regulatory tool. The implementation of this model of AI regulation by contract is significantly challenged by the lack of independence of the procurement function and the risks of bypassing of procurement-related constraints by the public sector entities using AI, their tech providers,

[390] A. Sanchez-Graells, *Public Procurement of Artificial Intelligence: Recent Developments and Remaining Challenges in EU Law,* Legal Tech Journal, 2024.

or both. It is further undermined by the growing gaps in the digital skills of many public institutions and public buyers. As a result, the risks identified remain largely unaddressed. Considering these conclusions, the EU should ensure that no digital technologies are procured or deployed in a way that infringes EU or domestic data and digital regulation, and that no contracting authority procures digital technologies without a sufficient impact assessment and without ensuring adequate digital skills to manage the process and the use of the technology throughout its lifecycle.

The best policy intervention to achieve these goals would involve a mix of:

- review of the EU procurement rules to create new mechanisms of assurance and, in particular, new mechanisms of impact assessment prior to the launch of a procedure for the procurement of digital technologies, which could be modelled on the impact assessments under the AI Act;
- creation of a single handbook of data and digital law applicable to procurement, on open access, curated and permanently updated by the European Commission;
- mandatory specific oversight mechanisms under independent authority to be developed at Member State level, to ensure compliance with the impact assessment and other regulatory requirements prior to the launch of the relevant procurement. Given that some Member States already have a similar function in place, a comparative study should be commissioned to extract the best practices and design features to be included in the EU requirement;
- coordination of the network of independent national authorities by the European Commission, perhaps through an expansion of the remit and a review of the institutional design of the future EU AI Office;
- mandatory specific tasks of digital professionalization and recruitment of digital skills into the procurement workforce, or a specialized body tasked with the

procurement of digital technologies and a credible plan for its continuous implementation and funding to be implemented by all Member States.

The implementation of such a policy mix would be more prescriptive than the strategy followed to date, whereby the facilitative approach taken by the European Commission in some areas (*e.g.*, procurement professionalization) has yielded limited results. The combined implementation of these measures would significantly minimize the risks and harms arising from inadequate processes of deployment of digital technologies by the public sector, as well as creating a level of protection of fundamental and individual rights, and of collective and social interest, that would be homogeneous across the EU[391].

2.4 The allocation of responsibility within the Product Liability Directive

As shown in the previous sections, neither the GDPR nor the AI Act appear to be effective measures to ensure that the public actor is not in information asymmetry and can have real control over the procurement process of private actors' AI systems used in public decision-making processes. Moreover, these regulations do not define a clear liability regime, which in the end has to be assessed between the parties during the negotiation process. Due to this framework, it is important to analyze the Product Liability Directive, which could open up to new and innovative solutions in assessing the liability in case of procurement by the public actor of AI systems provided for by private actors.

Indeed, independently of the AI Act, the Product Liability Directive suggests a general update of classical product liability, with a specific view, however, toward digital products more generally and AI more specifically. Indeed, legislative frameworks like the AI Act and liability regimes constitute two complementary approaches to regulating AI, directly (via specific regulation in the AI Act) and indirectly (via incentives generated by the liability

[391] *Ibid*.

framework)[392]. As presented in section 2.3.1, the AI Act outlines a regulatory and oversight framework for AI systems, particularly those considered as high-risk. Building on it, the Product Liability Directive now seeks to integrate the AI Act into civil (product) liability while aligning this field with the new risks and realities of the digital economy. Crucially, the Directive and the AI Act complement one another insofar as the latter does not contain any individual rights of affected persons, and the former lack specific, substantive rules on AI development and deployment[393].

2.4.1 General aim of the Directive and related concerns

The Product Liability Directive contains amendments for material as well as procedural product liability law. It seeks to ensure effective compensation of injured persons and prevent liability gaps arising from the peculiarities of AI technology. Essentially, and convincingly, victims of AI systems should enjoy the same level of protection as injured persons in scenarios not involving AI, including the ease of compensation. Second, this regime is supposed to foster legal certainty in an area of law plagued by hitherto unforeseeable *ad hoc* adaptation to AI systems. This, in turn, should facilitate the insurability of liability risks for companies and spur AI uptake and innovation, particularly with SMEs. The Commission notes that in its survey of 2020, liability constituted the top external barrier to AI adoption by companies in the EU (43%)[394]. Third, the incentivization of trustworthy AI has been a long-standing goal of the EU initiative to regulate AI systems[395]. Therefore, it is not surprising that the Directive is meant to facilitate the rollout

[392] H. Zech, *Liability for AI: public policy considerations*, ERA Forum, Journal of the Academy of European Law, 2021, p. 147 – 150.
[393] P. Hacker, *The European AI liability directives – Critique of a half-hearted approach and lessons for the future*, Computer Law & Security Review, 51, 2023.
[394] Ipsos, *European enterprise survey on the use of technologies based on AI*, 2020, p. 58.
[395] On this matter, please see Independent High-Level Expert Group on Artificial Intelligence, *Policy and Investment Recommendations for Trustworthy AI*, (April 8th, 2019); Recital 7 of the AI Act; J. Laux – S. Wachter – B. Mittelstadt, *Trustworthy Artificial Intelligence and the European Union AI Act: On the Conflation of Trustworthiness and the Acceptability of Risk*, 2022; M. Taddeo, *Modelling trust in artificial agents, a first*

and uptake of trustworthy AI. The Product Liability Directive, together with several other EU initiatives, also aims to 'minimize' the risks of digital products[396].

Moreover, the Product Liability Directive applies to physical products as well as software, including AI systems. It harmonizes the supposedly strict liability of manufacturers, and of other entities in the supply chain under certain conditions, based on product liability rooted in EU law. Furthermore, eligible damage positions under the Product Liability Directive are restricted to life, health, property, and loss of data, while Member State law is generally more open and may include the infringement of fundamental rights or primary financial loss. Finally, the Product Liability Directive, under Art. 3, provides for full harmonization.

2.5 Conclusions

Ultimately, governments should give greater attention to how they design and structure their contracts for services to develop and operate AI tools. Using contracting as a tool for algorithmic governance can allow governments and society to benefit from the improvements that AI tools can offer, while also helping to ensure that these tools will be designed and deployed responsibly. Citizens deserve to know about the algorithms that affect their lives and interests, and public knowledge about algorithms – or even just the potential for litigation seeking to review an algorithm – can itself provide some constraint on ill-considered and unfair algorithmic practices. After all, if AI vendors can operate in total secrecy, protected by trade secrets, and never expected to meet basic standards for

step toward the analysis of e-trust, Minds and Machines, 2010, p. 243; W. Pieters, *Explanation and trust: what to tell the user in security and AI?*, Ethics and Information Technology, 2011, p. 53; F. S. Grodzinsky - K. W. Miller – M. J. Wolf, *Developing artificial agents worthy of trust: 'Would you buy a used car from this artificial agent?'*, Ethics and Information Technology, 2011, p. 17-21; A. Ferrario – M. Loi – E. Viganò, *In AI we trust incrementally: A multi-layer model of trust to analyze human-artificial intelligence interactions*, Philosophy & Technology, 2020, p. 523.

[396] P. Hacker, *The European AI liability directives – Critique of a half-hearted approach and lessons for the future*, Computer Law & Security Review, 51, 2023.

data protection, privacy, algorithmic fairness, or public participation, then algorithmic accountability is nothing but a myth.

For this reason, procurement officers and the government officials they serve need to ensure that proper contractual duties and restraints are imposed on vendors' development and deployment of algorithmic tools. As they move toward automating a wider array of consequential governmental functions, governments would do well to spend additional time and thought during the procurement process to consider the ramifications of new AI tools and their intended uses, as well as their potentially unintended consequences. Indeed, as governments come to rely increasingly on AI tools to support tasks and functions that affect individuals and organizations, the demand for algorithmic accountability will only grow. Important concerns have already emerged about the fairness and transparency of AI technologies used by some governmental authorities, and it is evident that these concerns can be exacerbated when they contract out for the design, testing, and operation of AI tools. Private contractors may possess the analytic capacity that governments need for developing and running AI systems, but private companies' connections with the public will surely be more attenuated than will public actor's and its motives may not align as well with the delivery of public value. Private vendors also tend to prefer to conduct their work with less oversight and disclosure, often claiming trade secret protection over their algorithmic tools. Nevertheless, government contracting itself can operate as an important, tractable governance strategy.

Ultimately, to use AI tools responsibly, governments should seek to contract responsibly for the support and technology need to create such tools. Specifically, public officials and procurement officers should attend to four key issues. First, they should guarantee that AI contracts are drafted to ensure sufficient public transparency and to prevent vendors from claiming trade secret protection over all of their work. Second, government contracts

should obligate AI vendors to follow accepted privacy and security protocols – and to allow the government to access information needed to ensure those protocols are followed. Third, agencies should consider negotiating contracts that include substantive standards for responsible AI and insist that vendors follow procedures, such as periodic audits, to document their compliance with such standards. Finally, whenever agencies anticipate the need for public participation to inform the design and operation of AI tools, their AI contracts should obligate private vendors to cooperate in the process of public engagement. Contracting for algorithmic accountability is an immediately feasible strategy for governing a rapidly evolving and highly varied set of technological innovations. Government contracts can be designed and adapted so that they address the important, practical needs that lead governments to develop AI tools, while also respecting society's desire for public accountability and engagement. By using the procurement process to achieve greater algorithmic accountability, public officials can help provide a path toward a future in which AI is deployed responsibly to improve governmental performance[397].

3. Information sharing between private and public actors: a new data sharing model

In addition, the relationship between the public and private actors in using AI systems thanks to the creation of a mechanism of data sharing should be analyzed. As the aim is to maximize the potential of using and reusing data by sharing it with other entities, complex issues should be analyzed.

From a policy perspective, sharing data can be approached in various ways along a continuum, ranging from the most constricted, such as running analyses on the data and only sharing condensed bits of information (*e.g.*, aggregate or summary measures or

[397] C. Coglianese – E. Lampmann, *Contracting for Algorithmic Accountability*, All Faculty Scholarship, 2021, p. 2311.

model parameters), to the least constricted, such as providing full access as open data, for anyone to use at any time and for any application. In this context, the public actor may judge, based on its goals and permissions, which data share with whom and in what way. For example, achieving publics goals usually includes sharing data as openly as possible, by that maximizing the potential for use or reuse of that data for the benefit of all[398]. On the other hand, private companies usually refrain from sharing their own business-related data, as the reuse of such information by competing companies might negatively impact their own economic success[399].

In this field, in 2019 the Independent High-Level Expert Group on Artificial Intelligence in Policy and Investment Recommendations for Trustworthy AI[400] has stated that:

> 'A safe, secure and high-quality data infrastructure would enable Europe to better develop and train AI systems, which in turn can be steered towards applications that can facilitate the Sustainable Development Goals. A clear distinction must however be ensured between personal and nonpersonal data. A fundamental rights-based personal data infrastructure as put forward in the GDPR should be fostered and its enforcement should be ensured. At the same time, the sharing of industrial data still poses a significant challenge, creating obstacles to collaboration between organisations and, in some cases, stifling innovation. Therefore, industrial and research solutions must be found for processing and sharing data in a secure and privacy-respecting way, allowing everyone to reap its benefits to the fullest'.

[398] B. Mons – C. Neylon – J. Velterop – M. Dumontier - L. O. B. Da Silva Santos – M. D. Wilkinson, *Cloudy, increasingly FAIR; revisiting the FAIR data guiding principles for the European Open Science Cloud*. Information Services & Use, 2017, p. 49-56

[399] M. Tajabadi – L. Grabenhenrich – A. Ribeiro – M. Leyer – D. Heider, *Sharing Data With Shared Benefits: Artificial Intelligence Perspective,* Journal of Medical Internet Research, 2023.

[400] Independent High-Level Expert Group on Artificial Intelligence, *Policy and Investment Recommendations for Trustworthy AI*, (April 8th, 2019).

On this line, the European Commission, in its Communication entitled European Strategy for Data[401] has adopted a new fortified approach towards a regulation of personal and non-personal data. The starting point can be found in the awareness that 'Over the last few years, digital technologies have transformed the economy and society, affecting all sectors of activity and the daily lives of all Europeans (…)'[402]. In this scenario, it was clearly stated that: 'Data is at the centre of this transformation and more is to come. Data-driven innovation will bring enormous benefits for citizens, for example through improved personalized medicine, new mobility and through its contribution to the European Green Deal'[403]. On the basis of such premises, the European Commission significantly argues that: 'In a society where individuals will generate ever-increasing amounts of data, the way in which the data are collected and used must place the interests of the individual first, in accordance with European values, fundamental rights and rules'[404].

However, this anthropocentric vision also meets the needs of the European single market, in a multiple perspective typical of the European approach: together with the celebration of the individual protection of persons' fundamental rights and freedoms, statements concerning the opportunities of relevant social and economic development can be found[405]. In this direction, the Commission has stated that:

> 'Citizens will trust and embrace data-driven innovations only if they are confident that any personal data sharing in the EU will be subject to full compliance with the EU's strict data protection rules. At the same time, the increasing volume of non-personal industrial data and public data in Europe, combined with technological change in how

[401] European Commission, *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*, (February 19th, 2020).
[402] *Ibid*.
[403] *Ibid*.
[404] *Ibid*.
[405] F. Bravo, *Data Governance Act and Re-Use of Data in the Public Sector*, European Review of Digital Administration & Law, 2023.

the data is stored and processed, will constitute a potential source of growth and innovation that should be tapped'[406].

The enormous significance attributed to the processing of personal and non-personal data can be perfectly understood. Data are openly considered 'the new oil', not without some negative implications which need to be addressed, especially in the field of data protection law, competition law and AI law. However, the main value of data should not be founded in their direct economic worth, but in the set of capabilities that can be derived from themselves, by means of an accurate analysis. That is precisely the crux of the matter.

The great value of data mainly consists in supporting decision-making. Data and data analysis allow for better decisions, with huge benefits for natural and legal persons, such as citizens, associations, foundations, non-governmental organizations, enterprises and companies, and public administrations[407].

According to the above-mentioned Communication:

> 'Citizens should be empowered to make better decisions based on insights gleaned from non-personal data. And that data should be available to all – whether public or private, big or small, start-up or giant. This will help society to get the most out of innovation and competition and ensure that everyone benefits from a digital dividend. This digital Europe should reflect the best of Europe - open, fair, diverse, democratic, and confident'[408].

The approach adopted by the European Commission seems not to be the one based on the 'commodification' of personal and nonpersonal data, in order to have a monetary gain

---

[406] European Commission, *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*, (February 19th, 2020).

[407] F. Bravo, *Data Governance Act and Re-Use of Data in the Public Sector*, European Review of Digital Administration & Law, 2023.

[408] European Commission, *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*, (February 19th, 2020).

in the digital market of data, but the one that consider data as means of innovation and development for society, institutions and markets, both in private and public sector, 'to enable the EU to become the most attractive, most secure and most dynamic data-agile economy in the world – empowering Europe with data to improve decisions and better the lives of all of its citizens'[409].

Indeed, the EU aims to build a different model, in which data do not consist in 'commodities' or 'goods', but, first of all, in 'a value' available to all, as a key factor of growth, wealth and development, for the entire society, including citizens, public administrations, enterprises and other public and private bodies. In this direction the European Commission strongly specified – in its Communication on 'The European Strategy for Data' – that:

> 'The EU can become a leading role model for a society empowered by data to make better decisions – in business and the public sector. To fulfil this ambition, the EU can build on a strong legal framework – in terms of data protection, fundamental rights, safety and cybersecurity – and its internal market with competitive companies of all sizes and varied industrial base. If the EU is to acquire a leading role in the data economy, it has to act now and tackle, in a concerted manner, issues ranging from connectivity to processing and storage of data, computing power and cybersecurity. Moreover, it will have to improve its governance structures for handling data and to increase its pools of quality data available for use and re-use. Ultimately, Europe aims to capture the benefits of better use of data, including greater productivity and competitive markets, but also improvements in health and well-being, environment, transparent governance and convenient public services'[410,411].

[409] *Ibid*.
[410] *Ibid*.
[411] F. Bravo, *Data Governance Act and Re-Use of Data in the Public Sector*, European Review of Digital Administration & Law, 2023.

3.1 The AI Act: a missed opportunity?

Before analyzing the specific regulations provided for data sharing models, it is useful to understand whether in the AI Act the European legislator has regulated the matter *ad hoc*, considering also that an inherent close connection exists between the development of AI and data sharing models.

On this matter, Recital 68 of the AI Act states that:

> 'For the development and assessment of high-risk AI systems, certain actors, such as providers, notified bodies and other relevant entities, such as European Digital Innovation Hubs, testing experimentation facilities and researchers, should be able to access and use high-quality data sets within the fields of activities of those actors which are related to this Regulation. European common data spaces established by the Commission and the facilitation of data sharing between businesses and with government in the public interest will be instrumental to provide trustful, accountable and non-discriminatory access to high-quality data for the training, validation and testing of AI systems. For example, in health, the European health data space will facilitate non-discriminatory access to health data and the training of AI algorithms on those data sets, in a privacy-preserving, secure, timely, transparent and trustworthy manner, and with an appropriate institutional governance. Relevant competent authorities, including sectoral ones, providing or supporting the access to data may also support the provision of high-quality data for the training, validation and testing of AI systems'.

As shown, the AI Act in its recitals takes into account the issue of data sharing between businesses and governments in the public interest. However, Art. 59 severely limits the possibility of using data and information for purposes other than those for which they were

collected even for public purposes. Indeed, in the AI regulatory sandbox[412], personal data lawfully collected for other purposes may be processed solely for the purpose of developing, training and testing certain AI systems in the sandbox when all of the following conditions are met:

a. AI systems shall be developed for safeguarding substantial public interest by a public authority or another natural or legal person and in one or more of the following areas: (i) public safety and public health, including disease detection, diagnosis prevention, control and treatment and improvement of health care systems; (ii) a high level of protection and improvement of the quality of the environment, protection of biodiversity, protection against pollution, green transition measures, climate change mitigation and adaptation measures; (iii) energy sustainability; (iv) safety and resilience of transport systems and mobility, critical infrastructure and networks; (v) efficiency and quality of public administration and public services;
b. the data processed are necessary for complying with one or more of the requirements referred to in Chapter III, Section 2 where those requirements cannot effectively be fulfilled by processing anonymized, synthetic or other non-personal data;
c. there are effective monitoring mechanisms to identify if any high risks to the rights and freedoms of the data subjects, as referred to in Art. 35 of the GDPR and in Art. 39 of the EUDPR, may arise during the sandbox experimentation, as well as response mechanisms to promptly mitigate those risks and, where necessary, stop the processing;

[412] Under Art. 3, no. 3, of the AI Act, 'AI regulatory sandbox' means a controlled framework set up by a competent authority which offers providers or prospective providers of AI systems the possibility to develop, train, validate and test, where appropriate in real-world conditions, an innovative AI system, pursuant to a sandbox plan for a limited time under regulatory supervision.

d. any personal data to be processed in the context of the sandbox are in a functionally separate, isolated and protected data processing environment under the control of the prospective provider and only authorized persons have access to those data;
e. providers can further share the originally collected data only in accordance with Union data protection law; any personal data created in the sandbox cannot be shared outside the sandbox;
f. any processing of personal data in the context of the sandbox neither leads to measures or decisions affecting the data subjects nor does it affect the application of their rights laid down in Union law on the protection of personal data;
g. any personal data processed in the context of the sandbox is protected by means of appropriate technical and organizational measures and deleted once the participation in the sandbox has terminated or the personal data has reached the end of its retention period;
h. the logs of the processing of personal data in the context of the sandbox are kept for the duration of the participation in the sandbox, unless provided otherwise by Union or national law;
i. a complete and detailed description of the process and rationale behind the training, testing and validation of the AI system is kept together with the testing results as part of the technical documentation referred to in Annex IV;
j. a short summary of the AI project developed in the sandbox, its objectives and expected results are published on the website of the competent authorities; this obligation shall not cover sensitive operational data in relation to the activities of law enforcement, border control, immigration or asylum authorities.

In the light of the above, it appears that the possibility of setting up data sharing mechanisms that would allow the public actor to use data collected by private actors is extremely limited and, in any case, complex, given the very large number of conditions that the public actor would have to comply with.

3.2 Other limits of the Business to Government (B2G) data sharing

Additionally, it should be considered that there is currently not enough private sector data available for use by the public sector to improve evidence-driven policy-making and public services[413].

As such, the recommendations of an Expert Group[414] created by the Commission include the creation of national structures for Business to Government (B2G) data sharing, the development of appropriate incentives to create a data-sharing culture, and the suggestion to explore an EU regulatory framework to govern the public sector's re-use for the public interest of privately held data[415].

Indeed, due to their spontaneous nature and limited scale, B2G data-sharing collaborations are not yet sufficiently visible, transparent, nor are they scalable and easily repeatable processes. Public authorities, private companies, organizations and the general public are not always fully aware of the benefits of B2G data sharing. Public authorities lack clarity on what data is available and what it entails to create value from it. As a result, private companies, organizations and the general public might be less willing to share data that could otherwise be used to tackle societal challenges. Some ethical considerations also prevent responsible data sharing from occurring, in particular in

[413] A. Owusu, *Data sharing in the personal data economy. Does sharing mean caring?*, European Journal of Privacy Law & Technologies, 2023.
[414] For more information, see here: https://ec.europa.eu/digital-single-market/news-redirect/666643.
[415] European Commission, *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*, (February 19th, 2020).

relation to the modalities of transparent data collection, use and reuse. In particular, by maximizing the accessibility and use of data, data sharing inherently raises major concerns, such as those epitomized by the risk of reidentification and the consequences flowing from data-protection legislation. Moreover, it should be considered that, from an operational and technical perspective, the availability of trusted technical systems that enable 'safe' B2G data sharing is currently limited. From the data provider's perspective, sharing data can entail a variety of risks, from the disclosure of sensitive commercial information, trade secrets, customers' personal information or the reidentification of a customer in breach of that customer's right to privacy. Data anonymization, pseudonymization and aggregation are a few of the most common techniques to protect personal data, yet their proper application is far from trivial. In addition, the limited trust existing between a given private company or civil-society organization and the public-sector body at the time of the storage, access and processing of data further prevents those collaborations from happening. These operational and technical challenges also currently contribute to preventing B2G data-sharing collaborations from scaling into a thriving and sustainable ecosystem[416].

As such, the desire of the European legislator to favor the development of fruitful opportunities for cooperation between public and private actors should be recognized, especially for the purpose of using personal data and not to ensure the development of new technologies such as AI. However, on the other hand, this trend seems to contradict the approach taken so far by the European legislator, which aims both to protect the fundamental rights of citizens and to slow down the rise of the private actor in favor of the public ones.

[416] European commission, *Towards a European strategy on business-to-government data sharing for the public interest*, 2020.

Considering all the presented issues, the EU has recently proposed and enacted different laws aiming at fostering collaboration between the public and the private actor, as well as at enhancing transparency between them. Indeed, the promotion of AI-driven innovation is closely linked to the Data Governance Act[417] (DGA) (section 3.3), Data Act[418] (section 3.4) and the Open Data Directive[419] (ODD) (section 3.5), which have established trusted mechanisms and services for the re-use, sharing and pooling of data that are essential for the development of data-driven AI models of high quality.

3.3 Data Governance Act

The DGA aims 'to improve the conditions for data sharing in the internal market, by creating a harmonized framework for data exchanges and laying down certain basic requirements for data governance, paying specific attention to facilitating cooperation between Member States'[420] with the ultimate goal of further developing 'the borderless digital internal market and a human-centric, trustworthy and secure data society and economy'[421]. In addition to regulating data intermediaries (Arts. 10 – 15), the DGA also addresses the re-use by the private actors of data held by the public sector (Arts. 3 – 9), and it further contains special rules for organizations that require and use personal data for altruistic purposes (Arts. 16 – 25)[422].

In the context of this work, its analysis is of utmost importance as it testifies the new direction taken by the European legislator with respect to the need for free and fluid

[417] Regulation (EU) 2022/868 of the European Parliament and of the Council of 30 May 2022 on European data governance and amending Regulation (EU) 2018/1724 (Data Governance Act).
[418] Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonized rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act).
[419] Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information (Open Data Directive).
[420] Recital 3 of the DGA.
[421] *Ibid*.
[422] A. Owusu, *Data sharing in the personal data economy. Does sharing mean caring?*, European Journal of Privacy Law & Technologies, 2023.

sharing of personal data and information from the public actor to the private ones that tries to go beyond the strict boundaries already drawn by the GDPR. This is also with a view to strengthening public-private partnership initiatives that ensure greater transparency in the relationship between the parties.

Particularly, it is clear that the EU's focus has long been on data protection. This focus culminated to the enactment of the GDPR in 2016. However, in recent years, the focus has switched toward facilitating data economy and data sharing in Europe. To some extent, the data economy focus was already present during the policy-making of the GDPR. For many politicians and stakeholders, the regulation had the twin goals of protecting personal data and facilitating the free flow of personal data across the internal market[423]. Data reuse was also a hot topic during the negotiations. Another point is that to a certain degree the data economy focus was also explicitly embedded into the GDPR, which, according to Art. 20, gives data subjects a new right for data portability between different data controllers. While the idea was to facilitate data sharing and interoperability, the new portability right turned out to be problematic in many ways, particularly with respect to data reuse. Specifically, it is, therefore, possible to argue that the use of personal data, especially in the context of AI technologies, has been profoundly constrained in Europe by such regulatory restrictions.

In this context, it appears that the European legislator has reversed course, promoting, instead, the reuse of even personal data precisely through the DGA, acknowledging that 'It is important to enable a competitive environment for data sharing'[424]. Indeed, the DGA seeks to facilitate further sharing of personal data by introducing a concept of data altruism. Another core tenet in the new regulation is the reuse of data held by public sector

[423] P. D. König, *Fortress Europe 4.0? An Analysis of EU Data Governance Through the Lens of the Resource Regime Concept*, European Policy Analysis, 2022, p. 484–504.
[424] Recital 33 of the DGA.

bodies.

In this field, the goals of the DGA are ambitious. The primary goal is to facilitate data economy in Europe and improve the EU's digital single market. Emphasis is placed upon the public sector for which the planned data reuse and data sharing provide new material to innovate in AI and digital applications. Moreover, fairness, data protection, and lawfulness receive considerable attention in the DGA[425].

According to Art. 1, the DGA lays down (1) the conditions for reuse of data held by European public sector bodies; (2) a notification and supervisory framework for the provision of data intermediation services; (3) a framework for voluntary registration mechanism for entities that collect and process data made available on altruistic purposes; and (4) a framework for establishing a new European board for innovation in data economy.

Data altruism is defined in Art. 2 and it refers to the voluntary sharing of data on the basis of the consent of data subjects to process personal data pertaining to them, or permissions of data holders. This to allow the use of their non-personal data without seeking or receiving a reward that goes beyond compensation related to the costs that they incur where they make their data available for objectives of general interest as provided for in national law where applicable, such as healthcare, combating climate change, improving mobility, facilitating the development, production and dissemination of official statistics, improving the provision of public services, public policy-making or scientific research purposes in the general interest. In other words, data altruism is based either on the permission given by an organization for not-for-profit processing activities of non-personal data or the notion of consent in case personal data is involved.

The categories of data for reuse are defined in Art. 3. Accordingly, the DGA applies to data

[425] J. Ruohonen – S. Mickelsson, *Reflections on the Data Governance Act,* Digital Society, 2023.

held by public sector bodies which is protected on the grounds of commercial confidentiality, statistical confidentiality, protection of intellectual property, and protection of personal data. Thus, personal data held by public sector bodies is covered and hence also the GDPR applies.

There are also exclusions. The DGA does not cover data held by public undertakings, data held by public service broadcasters and their subsidiaries, data held by cultural establishments and educational institutions, data protected on the grounds of national security and defense, and data falling outside the scope of the public tasks of the public sector bodies concerned[426].

Moreover, the conditions for data reuse are defined in Art. 5. The general principles are nondiscrimination, transparency, proportionality, and proper justification without attempts to restrict competition. To ensure that data is properly protected, public sector bodies must ensure that personal data is anonymized, and commercially confidential data is properly modified, aggregated or otherwise handled with proper disclosure controls. Thus, the GDPR's concept of pseudonymization is not sufficient: proper anonymization is generally required for reuse of personal data. That said, Art. 5 provides also two alternative options: a secure processing environment controlled by a public sector body in case remote access is provided or reuse and processing at the physical premises of a public sector body. In all cases security must be guaranteed. This article also prohibits users of reused data from any attempts to re-identify data subjects.

Additionally, to help public sector bodies with their new tasks, Art. 7 specifies that the Member States are obligated to designate specific competent bodies. The support provided by these competent bodies includes technical guidance for data storage and data processing, help with anonymization, suppression, randomization, and other

[426] *Ibid*.

techniques that ensure privacy, confidentiality, integrity, and accessibility of personal data, state of the art privacy-preserving methods, deletion of commercially confidential information, support for consent and permission requests for reuse, and relevant contractual commitments.

The other important concept under Art. 16 is data altruism that the Member States are instructed to promote and facilitate. The DGA speaks about specific data altruism organizations, which are legal persons that operate on a not-for-profit basis without any dependencies on for-profit entities. As with data intermediation services, all data altruism organizations wanting to be officially recognized as data altruism organizations must be officially registered to public registries maintained by competent public sector bodies. These organizations must keep rigorous track of those processing data held by the organizations[427].

3.3.1 Possible concerns

The new DGA lays down frameworks for data reuse and data altruism under the supervision of competent public sector authorities. However, it seems problematic in that it relies on consent for the sharing and processing of personal data. Although in Europe consent builds upon the foundational concept of informational self-determination, the use of consent as a legal basis for processing personal data has long been criticized from empirical, legal, and technical perspectives. Indeed, consumers and users of digital applications and services do not really understand to what they are consenting to, even though the GDPR's Art. 4 explicitly states that 'consent means any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of

[427] *Ibid*.

personal data relating to him or her'. According to skeptical viewpoints, there is little reason to believe that things would be different for the noble goal of data altruism.

Similar points have been raised also regarding the reuse of public sector data for which the conditions for consent are often different. In other words, it is difficult to specify the purpose of processing personal data at the time of initial data collection in a context that involves further processing. To some extent, the European politicians and lawmakers seem to have been aware of these issues, given that Art. 25 in the DGA mentions the development of a specific European data altruism consent form. However, it can be challenging for data altruism consent to fully comply with the GDPR's consent requirements as reaching the full potential of data economy requires flexibility in processing activities[428].

Moreover, the DGA raises also other concerns about the personal data protection right and the GDPR. Three such concerns deserve a brief discussion.

First, the DGA seems to conflict with some of the fundamental principles of the GDPR. In particular, the GDPR's Art. 5 explicitly states that personal data should be only collected for 'specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes'. Although the same article specifies that this purpose limitation does not apply to public interest data archiving, scientific research, and statistical applications, the DGA's goal of public sector data reuse still raises a concern about whether personal data collected by public sector bodies will be used in a manner which is unexpected or risky to the data subjects. Given that the GDPR does not apply to anonymized data, the DGA's provision for data reuse under the GDPR's purpose limitation rests upon proper anonymization. The second concern follows. As is well-known, there are efficient algorithms for de-anonymization and re-identification of data subjects. The

[428] *Ibid*.

efficiency of such algorithms is likely to only increase with advances in machine learning and AI. Hence, it remains debatable how well the state-of-the-art privacy-preserving methods mentioned in the DGA can prevent de-anonymization and re-identification attempts. This concern applies equally to non-personal data held by public sector bodies under commercial confidentiality. The last concern is about national data protection authorities whose duties seem to substantially increase with the DGA. For instance, according to the DGA's Recital 15, prior to granting access for reuse of data, public sector bodies should carry out data protection impact assessments and consult data protection authorities in line with the GDPR's Art. 35 and 36. Such consultations cover also questions about anonymization. The DGA also mentions, in Recital 26, that the new competent bodies for monitoring intermediation services and data altruism organizations do not have a strict supervisory function, which is reserved for data protection authorities. Given the resources, coordination, and other problems already faced by European data protection authorities, a concern remains about how well the DGA will be administrated and enforced. In this field, the problems with the GDPR's enforcement provide an alarming precedent[429]. Finally, it will be interesting to understand how this information sharing mechanism, which aims to improve agreements between the public actor and private actors, can contribute to changing the existing balance of power by giving the public actor the ability to control what data and information certain private actors will use to make certain decisions through automated decision-making processes.

3.4 Data Act

As a premise, it should be noted that if the DGA is intended to promote the exchange of data and information between public and private actors, the Data Act, on the other hand,

[429] *Ibid*.

is entitled to regulate mostly the exchange between private actors and between private actors and consumers. In this context, although the scope of the Data Act may not seem particularly relevant to the subject of this work, its analysis becomes interesting for two reasons. Firstly, because Chapter V of the Data Act provides the discipline to making data available to public sector bodies, the Commission, the European Central Bank and Union bodies on the basis of an exceptional need. So residual cases are established for the exchange of data between the public and the private actor. Secondly, because the Data Act constitutes as an example of a change in the European legislator's strategy regarding the sharing of personal and non-personal data.

That being clarified, the Data Act has the ambitious goal to lay down 'harmonized rules on fair access to and use of data'[430] in the internal market. It is based on the perception of data as having an immense potential for the economy and society as a whole. Data are considered as non-rival resources that should thus be shared and reused in order for their potential to be unleashed.

From this perspective, data monopolization by private actors is considered insofar as an obstacle to the optimal allocation of data.

In addition to the concern for an economic efficient allocation of data, the Data Act is also driven by a certain idea of 'fairness'. Both private and public actors should be able to use and generate value from data that they have generated, which is considered as the characterization of data control. In this context, data monopolization by private actors is preventing governments and other businesses from obtaining and making use of data generated by them. It can also prevent third parties from accessing and using data. Therefore, the Data Act grants data access and use rights in different contexts and to

[430] Art. 1 of the Data Act.

different types of actors[431].

In this context, the policy vision has two-pillars, where both economic efficiency and a certain account of fairness[432] shall be covered. Particularly, the Data Act is based on the expectation that the prospect of deriving value from data will incentivize actors to act upon the data access and use rights, thus resulting in economic efficiency. Indeed, to reinforce the alignment with the economic efficiency goal, the Data Act coins a new concept of 'data literacy' that dedicated enforcement authorities, in their enforcement practices, shall promote towards the beneficiaries of rights. Not to be confused with digital literacy, data literacy is defined by the Data Act as the skills, knowledge and understanding that allows individuals and businesses to 'gain awareness of the potential value of the data they generate'[433] to incentivize them to take an active role in the data markets, viewed as empowerment[434].

Thus, similar to the DGA, the Data Act aims to provide the tools and to incentivize individuals and businesses to act as data market participants. While the DGA focuses on the institutions and structures that support and facilitate data sharing, the Data Act focuses on substantive rights on – or in relation to – data. As further discussed in this chapter, both legislations – and especially the Data Act – can thus be analyzed as property institutions, defined broadly and functionally as 'an institution for organizing the use of [data] as resources in society'[435, 436].

---

[431] C. Ducuing, *The Regulation of Data in the European Union: the Data Governance Act and the Data Act*, in (eds.) E. E. Akin – S. Klimbacher – G. Ziccardi, *Smart cities, artificial intelligence and digital transformation law*, Milano University Press, 2024.

[432] Fairness under the Data Act shall be understood in the sense of fair and undistorted (data) markets (Recital 6 of the Data Act).

[433] Recital 19 of the Data Act.

[434] C. Ducuing, *The Regulation of Data in the European Union: the Data Governance Act and the Data Act*, in (eds.) E. E. Akin – S. Klimbacher – G. Ziccardi, *Smart cities, artificial intelligence and digital transformation law*, Milano University Press, 2024.

[435] T.W. Merrill, *The Property Strategy*, University of Pennsylvania Law Review, 160, 2012, p. 35.

[436] C. Ducuing, *The Regulation of Data in the European Union: the Data Governance Act and the Data Act*, in (eds.) E. E. Akin – S. Klimbacher – G. Ziccardi, *Smart cities, artificial intelligence and digital transformation law*, Milano University Press, 2024.

In this context, Chapter II of the Data Act lays down data rights for the user enforceable against the ‘data holder’ (often the manufacturer), namely the actor having control over data. First, in principle, under Art. 3, connected products shall be designed in such a way that data are made accessible to the user, *i.e.*, by design compliance. Should that not be the case, according to Art. 4, then users shall dispose of a data access right enforceable against the data holder (*i.e.*, manufacturer). The data access right is explicitly aimed at enabling data use: data shall be provided with the relevant metadata, they shall be of the same quality as the ones available to the data holder and especially ‘machine-readable’. If technically feasible, they shall be provided continuously and in real-time. Second, according to Art. 5, users dispose of the right to require data holders to share data with a third party that they select. The selected third party and the data holder shall conclude a contract pertaining to the conditions for such data sharing, which can include a fair compensation – or in other words a fair price (Art. 8 and 9). Importantly, the user can exercise this right several times to the benefit of several third parties that they have selected and possibly with respect to the same data.

The Data Act also lays down provisions instrumental to the data access and portability rights, such as transparency requirements and the prohibition of manipulative or otherwise deceiving behaviors (known as ‘dark patterns’) impairing the autonomous exercise of their rights by users. However ambitious, these rights – data access and data portability – are as such not new. The novelty lies in the regulation of the use of data by the three categories of actors – data holder, user and selected third party – who find themselves in a triangular relationship with one another. In line with its property ambition, the Data Act seeks to assign the benefits of data use and value while mitigating the resulting harmful consequences on the other actors.

However, in walking this ridgeline, the Data Act is not without some contradictions and

unclarities, such as concerning the conditions under which the data holder may use data. In this respect, the Data Act distinguishes personal data – deemed to be exhaustively regulated by the GDPR – from non-personal data which the data holder shall use only based on a contract with the user (Art. 4(13)). This provision has been interpreted as an indirect exclusive allocation of connected product data to the user, literally contradicting the spirit of the Data Act that data use and value should be shared more broadly and especially between these actors that generate data – users and data holders for that sake. In contrast, the user shall in principle be free to use data for any lawful purpose. As to the selected third party, they shall process data for the purpose and conditions agreed with the user, which can be assimilated to a specific (*i.e.*, triangular) form of a data license. In addition, the Data Act lays down data use limitations to preserve the rights and legitimate interests of the other actors in the triangle. For example, trade secrets of the data holder shall be preserved by the user and selected third party. Another example is the prohibition to all actors in the triangle to use data to draw insights into the economic situation, assets and production methods of the other actors in the triangle[437].

Moreover, it should be noted that the Data Act is based on the observation that, in the absence of a property legal status of data, most data access and use across the EU actually relies on and is regulated through contracts. However, in the presence of power imbalance between businesses, influential actors may easily reserve data use and value for themselves or share data only under prohibitive unfair contractual terms to the detriment of weaker and smaller businesses (especially SMEs). The EU already disposes of a comprehensive set of rules protecting consumers against unfair commercial

[437] *Ibid*.

practices[438] and contractual terms of traders[439,440]. These rules apply horizontally, namely in principle to all sectors and irrespective of the nature of the goods and services. In contrast, with B2B, the EU has proceeded with a piecemeal approach that addresses specific problems in specific sectors. Following this direction, Chapter IV of the Data Act continues this piecemeal approach, with its entry point being this time the traded object, namely data, irrespective of the sector. Moreover, Chapter IV of the Data Act regulates contractual terms between businesses that concern the access to and use of data or the liability and remedies for the breach or the termination of data-related obligations. On the twofold condition that a term has been unilaterally imposed by a business onto the other and that such term is unfair, it shall not be binding onto that other business (Art. 13(1)).

Additionally, within the context of the B2G data sharing obligations, the Data Act aims to find a balance between the pursuit of public or general interests and the preservation of the freedom to conduct a business, with the high ambition to come up with a horizontal legal regime, namely, to a significant extent, a purpose and sector agnostic one. The principle is that public sector bodies, the Commission, the European Central Bank or a Union body request a business to share data only in case of an 'exceptional need'. An exceptional need can, first, be identified where the requested data – whether personal or non-personal – are necessary to respond to a 'public emergency', namely an exceptional situation, limited in time, that negatively affects the population of the Union or the whole or part of a member State, with a risk of serious and lasting repercussions for, among

[438] See the Directive 2005/29/EC of the European Parliament and of the Council of 11 May 2005 concerning unfair business-to-consumer commercial practices in the internal market and amending Council Directive 84/450/EEC, Directives 97/7/EC, 98/27/EC and 2002/65/EC of the European Parliament and of the Council and Regulation (EC) No 2006/2004 of the European Parliament and of the Council (Unfair Commercial Practices Directive).

[439] See the Council Directive 93/13/EEC of 5 April 1993 on unfair terms in consumer contracts.

[440] See the Directive 2011/83/EU of the European Parliament and of the Council of 25 October 2011 on consumer rights, amending Council Directive 93/13/EEC and Directive 1999/44/EC of the European Parliament and of the Council and repealing Council Directive 85/577/EEC and Directive 97/7/EC of the European Parliament and of the Council Text with EEA relevance.

other things, living conditions or economic or financial stability (Art. 2(29)). Then, the condition is that there is no timely and effective manner, under equivalent conditions, for public authorities to obtain the necessary data. Second, in contrast to responding to a public emergency, the other circumstances in which public authorities may request data do not address the substantive nature of the 'exceptional need'. Also, and in contrast to the original Commission proposal, they concern only non-personal data in the face of the serious risk of interference of Chapter V of the Data Act with data protection. Public authorities may then request data from businesses where the absence of such data would prevent them from fulfilling a task carried out in the public interest and based on EU or national law. Finally, public authorities may request data from businesses, irrespective of the purpose, when they have exhausted all the other means at their disposal to obtain data in a timely manner. This reads as a catch-all provision for all sorts of situations that cannot be anticipated in advance. This implies that they could neither purchase such data on data markets at market price, nor rely on existing or future data sharing obligations.

Furthermore, Chapter V of the Data Act regulates the conditions in which public authorities request data from businesses and the conditions in which they may use data, to prevent abuse on their behalf and to minimize the detrimental impact on businesses (Art. 17 to 19). In particular, public authorities shall 'not use data in a manner incompatible with the purpose' for which they were requested, a wording reminiscent of the purpose limitation principle under the GDPR[441]. They shall cease to use and erase data as soon as no longer necessary for such purpose and they shall implement measures to preserve the confidentiality and integrity of data, such requirements being reminiscent of GDPR principles[442].

The possibility to share data with third parties is strictly regulated, with a more lenient

[441] Art. 5(1)(b) and Art. 6(4) of the GDPR.
[442] Art. 5(1)(c) and (f) of the GDPR.

approach where the exceptional need implies that research organizations also process data (Art. 21). Public authorities are under the same prohibition concerning the use of data to derive insights about the economic situation, assets and production or operation methods of the business who shared data. Disclosure of trade secrets as a result of mandatory data sharing shall be requested only when necessary and subject to the adoption of protective measures by public authorities.

Lastly, a debated question has been whether mandatory data sharing shall be subject to compensation, to which the Data Act offers a differentiated answer depending on the legal basis. When based on the exceptional need taken from responding to a public emergency, no compensation can be claimed by the business, allegedly justified by the magnitude of such exceptional need. In the other cases, businesses may claim fair compensation, namely one that is limited to the costs of making data available in addition to a reasonable margin (Art. 20)[443].

3.5 Open Data Directive

Finally, in order to provide a comprehensive view of the European regulatory framework regarding the data sharing models, the ODD is analyzed.

Replacing the PSI Directive[444], the ODD was enacted in 2019 in order to facilitate and regulate access to data held by the public sector. Under Art. 1:

> 'In order to promote the use of open data and stimulate innovation in products and services, this Directive establishes a set of minimum rules governing the re-use and the practical arrangements for facilitating the re-use of: (a) existing documents held by

[443] C. Ducuing, *The Regulation of Data in the European Union: the Data Governance Act and the Data Act*, in (eds.) E. E. Akin – S. Klimbacher – G. Ziccardi, *Smart cities, artificial intelligence and digital transformation law*, Milano University Press, 2024.

[444] Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information.

public sector bodies of the Member States; (b) existing documents held by public undertakings that are: (i) active in the areas defined in Directive 2014/25/EU; (ii) acting as public service operators pursuant to Article 2 of Regulation (EC) No 1370/2007; (iii) acting as air carriers fulfilling public service obligations pursuant to Article 16 of Regulation (EC) No 1008/2008; or (iv) acting as Community shipowners fulfilling public service obligations pursuant to Article 4 of Regulation (EEC) No 3577/92'.

As such, the ODD must be seen as part of a package of measures aiming to facilitate the creation of a common European data space[445], with the EU progressively asserting dominance over national data policy. Its rules apply to data held by public sector bodies, that are protected on grounds of commercial confidentiality, statistical confidentiality, protection of intellectual property rights of third parties, and the protection of personal data (Art. 3(1)). It applies solely to non-personal data (Art. 1(2)(h)) and does not extend to documents protected by intellectual property rights of third parties (Art. 1(2)(c)) or documents qualified as confidential information[446]. More specifically, the provisions solely apply to data which is already accessible on the basis of existing European and national access regimes (Article 1(3)), new access rights as such are, therefore, not established. Moreover, the ODD does bring with it some notable new provisions on 'access', as it obliges public bodies to make dynamic data available by public bodies immediately after collection (Art. 5(5)), as long as this does not impose a disproportionate effort (Art. 5(6))[447]. Most striking, though, is the new chapter on high-value datasets (HVDs), which provides for a list of HVDs by way of an implementing act.

Additionally, one of the most notable changes include further limiting the possibilities for

[445] Key among instruments for achieving this purpose are the DGA and the Data Act, as presented in the previous sections.
[446] M. Leistner – L. Antoine, *IPR and the Use of Open Data and Data Sharing Initiatives by Public and Private Actors*, 2022. Particularly, this is a study commissioned by the European Parliament's Policy Department for Citizens' Rights and Constitutional Affairs at the request of the Committee on Legal Affairs in 2022.
[447] H. Broomfield, *Where Is Open Data in the Open Data Directive?*, Information Polity, 2023, p. 175 – 188.

charging for data, which constitutes a market barrier for SMEs and start-ups (Art. 6), strengthening the transparency requirements for PSI-related public-private agreements and limiting agreements that could lead to exclusive re-use by private partners (Art. 12). Research data resulting from public funding (Art. 10) and public undertakings, such as transport operators and energy providers, are now also included in its scope because these are deemed as having tremendous re-use potential.

Nevertheless, although open data has been elevated into the directive's title, replacing PSI in the short name and is recurrent in the preamble, it is rarely mentioned in the substantive provisions. The most notable is in Art. 1(1), which states that 'In order to promote the use of open data and stimulate innovation in products and services, this Directive establishes a set of minimum rules governing the re-use and the practical arrangements for facilitating the re-use'. However, this is essentially limited to the scope of the potential for open data. It establishes that the directive is an instrument to promote the re-use of open data rather than regulating open data per se. Otherwise, only two other mentions are made, here both in relation to HVDs, namely in Art. 14(1) stating that 'investments made by the Member States in open data approaches, such as investments into the development and roll-out of certain standards, shall be taken into account and balanced against the potential benefits from inclusion in the list', and in Art. 16, which provides for the establishment of the Committee on Open Data and the Re-use of Public Sector Information to assist in defining the measures for the HVD Implementing Act(s)[448].

3.6 Elements of consistency within EU data legislation

At first glance, this overview has shown that EU data legislation undoubtedly has horizontal – or universal – ambitions. This is visible with the DGA, which attempts to

[448] *Ibid*.

incentivize data sharing in a wide range of scenarios and in particular both market and non-market ones. The horizontality of EU data legislation is also visible with B2G mandatory data sharing that applies irrespective of the type of data, type of public authority and even, to a large extent, irrespective of the type of exceptional need at stake. Similarly, the regulation of connected product data applies irrespective of the sector. As a broader range of sectors and products become smart or connected, the scope will naturally broaden.

Finally, the ambition of EU data legislation to provide for a horizontal regulation of data is visible from the harmonized legal conditions for mandatory data sharing between businesses which applies irrespective of the sector, context and, even more striking, even rationale for mandatory data sharing. Even more important, a set of common features and principles can be identified. With the exception of the regulation of data processing services, EU data legislation is based on the same premise that, as non-rival resources, data shall, in principle, be shared and broadly used in order to optimize their economic potential particularly in the field of AI. Then, the whole of EU data legislation consists of balancing this principle with contextual specificities requiring more closedness. The legitimate rights and interests of different actors – and especially public actor and private ones – are symptomatically put increasingly on equal footing. On the one hand, they may tip the scales towards closedness, and, on the other hand, they are interpreted, especially under the DGA, as legal mechanisms likely to support data sharing akin to property entitlements.

This raises the question whether EU data legislation is compatible with data protection. In particular, it remains to be assessed whether data protection may continue as we know it while being subsumed, as it is, into such an economic and market framework. This can be associated with the property ambitions of EU data legislation that consistently aims to

provide for a form of initial allocation of data (Data Act) then allowing for secondary allocation through different types of data sharing (especially under the DGA). Together, the DGA and the Data Act are, therefore, forming a private law infrastructure for data. Data generation[449] constitutes a recurring justification under the Data Act for the granting of rights to use data or prevent their use by others, observable with the regulation of connected product data and with the data-specific regulation of unfair commercial practices between businesses. EU data legislation does not provide for a definition of activities deemed 'generative' of data, despite the fact that this notion produces legal effects. Subject to future case law, a broad concept of data generation seems to be emerging encompassing even the mere passive state of being - or having one's equipment – 'datafied'. Relatedly, however imprecise, the notion of 'data control', as supported by EU data legislation, appears to constitute a consistent principle, expected to serve as a bridge between several contradictory objectives: (i) while data cannot and shall not be exclusively 'owned', data control constitutes a soft property functional equivalence for data, likely to legally support different types of data sharing activities and to enable individuals and businesses to generate value from data; (ii) data control is also demonstrably expected to bridge private interests of individuals and businesses with respect to 'their' data with what the EU legislature considers as the general interest, namely unleashing the (economic) potential of data while protecting EU industrial interests on the international plane. In that sense, data control lies at the core of the neo-mercantilist project of the EU; and (iii) EU data legislation aims to bring, under the banner of data control, different types of rights and legitimate interests, and especially these of, respectively, individuals and businesses concerning, respectively, personal and non-personal data.

This reinforces the sense of a unified 'EU data law' for which, undoubtedly, data control

[449] Recital 6 of the Data Act.

constitutes the cornerstone[450].

[450] C. Ducuing, *The Regulation of Data in the European Union: the Data Governance Act and the Data Act*, in (eds.) E. E. Akin – S. Klimbacher – G. Ziccardi, *Smart cities, artificial intelligence and digital transformation law*, Milano University Press, 2024.

CHAPTER V: US APPROACH: THE USE OF AI BY THE PUBLIC ACTOR IN A COMPARATIVE PERSPECTIVE

Contents: 1. A comparative perspective: the approach of the United States; 2. The US regulation of the 'AI-driven public actor'; 2.1 Protection of citizens' rights within the Federal Government use of AI; 2.2 Public Procurement of AI; 2.3 Evaluations on the US approach used in the AI Executive Order; 4. Memorandum on Harnessing Artificial Intelligence to Fulfill National Security Objectives; 5. Comparing US and EU approaches in regulating AI; 5.1 The AI Executive Order and the AI Act in regulating the use of AI by the public actor; 6. Concluding Remarks.

1. A comparative perspective: the approach of the United States

Considering the analysis carried out in the previous chapters, it is clear the EU's attempt to regulate the use of AI by the public actor. While waiting to understand whether the AI Act, as well as all the other regulations analyzed, will have the same 'Brussels effect'[451] that the GDPR already had, it is interesting to examine, from a comparative perspective, which approach is used in other legal systems.

To this end, taking advantage of the publication of an *ad hoc* provision on the use of AI technologies by the United States (US), even though the text has been revoked by the current US Administration[452], this chapter analyses the previous act, also to understand

[451] A. Bradford, *The Brussels Effect: How the European Union Rules the World*, Faculty Books. 2020, p. 232. According to A. Bradford, the Brussels Effect shows how the EU has acquired such power, why multinational companies use EU standards as global standards, and why the EU's role as the world's regulator is likely to outlive its gradual economic decline, extending the EU's influence long into the future.

[452] Please note that on January 23rd, 2025 the current US Administration has revoked the AI Executive Order stating that '(a) The APST, the Special Advisor for AI and Crypto, and the APNSA shall immediately review, in coordination with the heads of all agencies as they deem relevant, all policies, directives, regulations, orders, and other actions taken pursuant to the revoked Executive Order 14110 of October 30, 2023 (Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence). The APST, the Special Advisor for AI and Crypto, and the APNSA shall, in coordination with the heads of relevant agencies, identify any actions taken pursuant to Executive Order 14110 that are or may be inconsistent with, or present obstacles to, the policy set forth in section 2 of this order. For any such agency actions identified, the heads of agencies shall, as appropriate and consistent with applicable law, suspend, revise, or rescind such actions, or propose suspending, revising, or rescinding such actions. If in any case such suspension, revision, or rescission cannot be finalized immediately, the APST and the heads of agencies shall promptly take steps to provide all available exemptions authorized by any such orders, rules, regulations, guidelines, or policies, as

whether the approach used can be considered more effective than the one adopted in the EU.

On 30 October 2023, the Executive Order[453] on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (AI Executive Order) has been adopted, marking a key development in the global race to regulate AI. The AI Executive Order articulates a vast array of commitments on a variety of sectors that are likely to be influenced by the implementation of AI systems on a large scale. Particularly, the US government seems to have adopted a more pragmatic approach, despite the predominantly programmatic nature of the AI Executive Order, which nevertheless embodies a clear agenda-setting ambition. The AI Executive Order mandates cover a variety of areas: safety and security, innovation and competition, equity and civil rights, workers' rights, protection of weak parties (students, patients, passengers, and consumers), privacy, as well as the promotion of the use of AI by the federal government. It then aims to strengthen the American leadership abroad.

Thus, the Executive Order intends to pave the way for the adoption of measures that would significantly impact different key sectors, establishing the necessary safeguards to protect individual and societal interests *vis-à-vis* the spread of AI systems. It aims to create conditions for the US to lead the AI revolution, by exploiting the full potential of technology without hindering human rights and freedoms.

The US approach to AI regulation relies on two primary strategies: establishing guidelines and standards through federal agencies and encouraging self-regulation within the industry. Additionally, it is guided by several overarching objectives, including fostering

appropriate and consistent with applicable law, until such action can be finalized. (b) Within 60 days of this order, the OMB Director, in coordination with the APST, shall revise OMB Memoranda M-24-10 and M-24-18 as necessary to make them consistent with the policy set forth in section 2 of this order'.

[453] Executive orders are presidential directives that do not rely on a specific constitutional provision. They result from the exercise of the executive power vested in the President, which has broad enforcement authority under Art. 2 of the Constitution.

openness and competitiveness in the AI-driven economy, enhancing safety while managing risks, and maintaining a technological edge over global competitors. These strategies and objectives have shaped the US approach to AI regulation, which is characterized as market-driven, sector-specific, and primarily 'vertically' focused. It is highly decentralized across federal agencies and grounded in a risk management framework[454].

Given the scope of the AI Executive Order (and more specifically, the existence of specific references to the use of AI by the federal government) as well as the general aim of this work, two aspects are analyzed in the following paragraphs: (i) the protection of citizens' civil rights, especially when AI is used by the public actor, (ii) the safeguards provided due to the use by the US public actor of 'private-provided AI technologies'.

2. The US regulation of the 'AI-driven public actor'

Following the analysis of the AI Executive Order, it is not surprising to note that several of its parts concern legal aspects related to the use of AI systems by the public actor. Indeed, it states that:

> 'It is important to manage the risks from the Federal Government's own use of AI and increase its internal capacity to regulate, govern, and support responsible use of AI to deliver better results for Americans. These efforts start with people, our Nation's greatest asset. My Administration will take steps to attract, retain, and develop public service-oriented AI professionals, including from underserved communities, across disciplines – including technology, policy, managerial, procurement, regulatory, ethical, governance, and legal fields – and ease AI professionals' path into the Federal Government to help harness and govern AI. The Federal Government will work to

[454] D. Tatevik, *The U.S. Approach to AI Regulation: Federal Laws, Policies, and Strategies Explained*, (September 09th, 2024).

ensure that all members of its workforce receive adequate training to understand the benefits, risks, and limitations of AI for their job functions, and to modernize Federal Government information technology infrastructure, remove bureaucratic obstacles, and ensure that safe and rights-respecting AI is adopted, deployed, and used'.

Section 10 then establishes specific measures to advance Federal Government use of AI. Indeed, Section 10.1(b) details a set of governance reforms to be implemented in view of the Director of the Office of Management and Budget (OMB)'s guidance to strengthen the effective and appropriate use of AI, advance AI innovation, and manage risks from AI in the Federal Government. Section 10.1(b) includes the following:

'The Director of OMB's guidance shall specify, to the extent appropriate and consistent with applicable law:

1) the requirement to designate at each agency within 60 days of the issuance of the guidance a Chief Artificial Intelligence Officer who shall hold primary responsibility in their agency, in coordination with other responsible officials, for coordinating their agency's use of AI, promoting AI innovation in their agency, managing risks from their agency's use of AI;
2) the Chief Artificial Intelligence Officers' roles, responsibilities, seniority, position, and reporting structures;
3) for [covered] agencies […], the creation of internal Artificial Intelligence Governance Boards, or other appropriate mechanisms, at each agency within 60 days of the issuance of the guidance to coordinate and govern AI issues through relevant senior leaders from across the agency;
4) required minimum risk-management practices for Government uses of AI that impact people's rights or safety, including, where appropriate, the following practices derived from Office of Science and Technology Policy (OSTP)'s

Blueprint for an AI Bill of Rights and the NIST AI Risk Management Framework: (i) conducting public consultation; (ii) assessing data quality; (iii) assessing and mitigating disparate impacts and algorithmic discrimination; (iv) providing notice of the use of AI; (v) continuously monitoring and evaluating deployed AI; (vi) and granting human consideration and remedies for adverse decisions made using AI;

5) specific Federal Government uses of AI that are presumed by default to impact rights or safety;
6) recommendations to agencies to reduce barriers to the responsible use of AI, including barriers related to information technology infrastructure, data, workforce, budgetary restrictions, and cybersecurity processes;
7) requirements that [covered] agencies […] develop AI strategies and pursue high-impact AI use cases;
8) in consultation with the Secretary of Commerce, the Secretary of Homeland Security, and the heads of other appropriate agencies as determined by the Director of OMB, recommendations to agencies regarding:
    a. external testing for AI, including AI red-teaming for generative AI, to be developed in coordination with the Cybersecurity and Infrastructure Security Agency;
    b. testing and safeguards against discriminatory, misleading, inflammatory, unsafe, or deceptive outputs, as well as against producing child sexual abuse material and against producing non-consensual intimate imagery of real individuals (including intimate digital depictions of the body or body parts of an identifiable individual), for generative AI;
    c. reasonable steps to watermark or otherwise label output from generative AI;

d. application of the mandatory minimum risk-management practices defined this section to procured AI;
e. independent evaluation of vendors' claims concerning both the effectiveness and risk mitigation of their AI offerings;
f. documentation and oversight of procured AI;
g. maximizing the value to agencies when relying on contractors to use and enrich Federal Government data for the purposes of AI development and operation;
h. provision of incentives for the continuous improvement of procured AI; and
i. training on AI in accordance with the principles set out in this order and in other references related to AI listed herein; and

9) requirements for public reporting on compliance with this guidance'.

As presented, Section 10.1(b) of the AI Executive Order establishes two sets or types of requirements. First, there are internal governance requirements implemented to protect citizens' rights, and these revolve around the appointment of Chief Artificial Intelligence Officers (CAIOs), AI Governance Boards, their roles, and support structures. This set of requirements seeks to strengthen the ability of Federal Agencies to understand AI and to provide effective safeguards in its governmental use also from a rights perspective[455] (section 2.1). Second, there are external governance requirements that revolve around the public actor's ability to control and challenge tech providers (section 2.2).

Therefore, in the following paragraphs the analysis of these two aspects is presented also considering the Memorandum on Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence (Memorandum M-24-10), which has been issued by the US government on March 28th, 2024 in light of the prescriptions of the

[455] *Some thoughts on the US' Executive Order on the Safe, Secure, and Trustworthy Development and Use of AI,* How to crack a nut, A blog on EU Economic Law, (November 7th, 2023).

AI Executive Order aiming to directs agencies to advance AI governance and innovation while managing risks from the use of AI in the Federal Government, particularly those affecting the rights and safety of the public.

2.1 Protection of citizens' rights within the Federal Government use of AI

From the analysis of the text of the AI Executive Order, it is possible to understand how, at least at the programmatic level, the US government has paid special attention to the protection of citizens' rights, especially in the context of the use of AI systems by the public actor. Indeed, the AI Executive Order states that:

> 'Artificial Intelligence policies must be consistent with my Administration's dedication to advancing equity and civil rights. My Administration cannot – and will not – tolerate the use of AI to disadvantage those who are already too often denied equal opportunity and justice. From hiring to housing to healthcare, we have seen what happens when AI use deepens discrimination and bias, rather than improving quality of life.
>
> […]
>
> Americans' privacy and civil liberties must be protected as AI continues advancing. Artificial Intelligence is making it easier to extract, re-identify, link, infer, and act on sensitive information about people's identities, locations, habits, and desires. Artificial Intelligence's capabilities in these areas can increase the risk that personal data could be exploited and exposed. To combat this risk, the Federal Government will ensure that the collection, use, and retention of data is lawful, is secure, and mitigates privacy and confidentiality risks. Agencies shall use available policy and technical tools, including privacy-enhancing technologies (PETs) where appropriate, to protect privacy and to combat the broader legal and societal risks – including the chilling of First

Amendment rights – that result from the improper collection and use of people's data.

[…]

It is important to manage the risks from the Federal Government's own use of AI and increase its internal capacity to regulate, govern, and support responsible use of AI to deliver better results for Americans. These efforts start with people, our Nation's greatest asset. My Administration will take steps to attract, retain, and develop public service-oriented AI professionals, including from underserved communities, across disciplines – including technology, policy, managerial, procurement, regulatory, ethical, governance, and legal fields – and ease AI professionals' path into the Federal Government to help harness and govern AI. The Federal Government will work to ensure that all members of its workforce receive adequate training to understand the benefits, risks, and limitations of AI for their job functions, and to modernize Federal Government information technology infrastructure, remove bureaucratic obstacles, and ensure that safe and rights-respecting AI is adopted, deployed, and used'.

Consistently, the Memorandum M-24-10 establishes new requirements and recommendations that, both independently and collectively, address the specific risks from relying on AI to inform or carry out agency decisions and actions, particularly when such reliance impacts the rights and safety of the public. To address these risks, it requires agencies to follow minimum practices when using safety-impacting AI and rights-impacting AI, and it enumerates specific categories of AI that are presumed to impact rights and safety.

As this brief introduction shows, in its programmatical statements and principles the AI Executive Order (as well as the Memorandum M-24-10) focuses primarily on the effect the use of AI can have on citizens' rights, as mandating a commitment from the American government to prevent the risks associated with the use of such systems from impacting

the exercise of those rights. Indeed, advancing equity and civil rights is considered a prominent aim, directing agencies to combat algorithmic discrimination. Furthermore, the AI Executive Order intends to provide guidance to federal agencies to keep AI algorithms from being used to exacerbate discrimination, as well as to address algorithmic discrimination through training, technical assistance, and coordination between the Department of Justice and Federal civil rights offices on best practices for investigating and prosecuting civil rights violations related to AI. Additionally, the AI Executive Order also calls on Congress to pass bipartisan data privacy legislation – one that would recognize privacy as a human right and give citizens back control of their personal data. This is especially crucial given the ongoing bulk (meta) data collection and mass surveillance by both governments and private actors[456]. These can be considered necessary steps in the right direction, as they fall short of what could reasonably be expected of governments adequately protecting people's privacy, especially in terms of informational self-determination and sovereignty and data minimization. Lastly, the AI Executive Order also addresses some ethical concerns with the government's use of AI, such as discrimination risks and 'unsafe decisions'. To ensure the responsible and secure government's use of AI, the AI Executive Order suggests issuing guidance for agencies' use of AI, including clear standards to protect rights and safety, improve AI procurement, and strengthen AI deployment. The AI Executive Order also wants to help agencies to acquire specified AI products and services faster, more cheaply, and more effectively through more rapid and efficient contracting, accelerating the rapid hiring of AI professionals, and providing AI training for employees – all of which are laudable endeavors[457].

[456] S. Zuboff, *The Age of Surveillance Capitalism*, Public Affairs, 2019.
[457] M. Wörsdörfer, *Biden's Executive Order on AI and the E.U.'s AI Act: A Comparative Computer-Ethical Analysis,* Philosophy & technology, 2024, p. 74.

In order to reach all these goals, the Memorandum M-24-10 requires:

1. to carry out a complete AI impact assessment. Specifically, agencies should update their impact assessments periodically and leverage them throughout the AI's lifecycle. In their impact assessments, agencies must document at least the following: (i) the intended purpose for the AI and its expected benefit, supported by specific metrics or qualitative analysis; (ii) the potential risks of using AI, as well as what, if any, additional mitigation measures, beyond these minimum practices, the agency will take to help reduce these risks; (iii) the quality and appropriateness of the relevant data;
2. to test AI for performance in a real-world context. Particularly, agencies must conduct adequate testing to ensure the AI, as well as components that rely on it, will work in its intended real-world context. Such testing should follow domain-specific best practices, when available, and should take into account both the specific technology used and feedback from human operators, reviewers, employees, and customers who use the service or are impacted by the system's outcomes. Testing conditions should mirror as closely as possible the conditions in which the AI will be deployed. Through test results, agencies should demonstrate that the AI will achieve its expected benefits and that associated risks will be sufficiently mitigated, or else the agency should not use the AI;
3. to carry out an independent evaluation of the AI system. Agencies, through the CAIO, an agency AI oversight board, or other appropriate agency office with existing test and evaluation responsibilities, must review relevant AI documentation to ensure that the system works appropriately and as intended, and that its expected benefits outweigh its potential risks;
4. to conduct ongoing monitoring. In addition to pre-deployment testing, agencies

must institute ongoing procedures to monitor degradation of the AI's functionality and to detect changes in the AI's impact on rights and safety;

5. to evaluate regularly risks from the use of AI. The monitoring process must include periodic human reviews to determine whether the deployment context, risks, benefits, and agency needs have evolved. Agencies must also determine whether the current implementation of the memorandum's minimum practices adequately mitigates new and existing risks, or whether updated risk response options are required;
6. to mitigate emerging risks to rights and safety. Upon identifying new or significantly altered risks to rights or safety through ongoing monitoring, periodic review, or other mechanisms, agencies must take steps to mitigate those risks, including, as appropriate, through updating the AI to reduce its risks or implementing procedural or manual mitigations, such as more stringent human intervention requirements;
7. to ensure adequate human training and assessment. Agencies must ensure there is sufficient training, assessment, and oversight for operators of the AI to interpret and act on the AI's output, combat any human-machine teaming issues (such as automation bias), and ensure the human-based components of the system effectively manage risks from the use of AI;
8. to provide additional human oversight, intervention, and accountability as part of decisions or actions that could result in a significant impact on rights or safety. Agencies must assess their rights-impacting and safety-impacting uses of AI to identify any decisions or actions in which the AI is not permitted to act without additional human oversight, intervention, and accountability. When immediate human intervention is not practicable for such an action or decision, agencies must ensure that the AI functionality has an appropriate fail-safe that minimizes the risk

of significant harm;

9. to provide public notice and plain-language documentation. Agencies must ensure, to the extent consistent with applicable law and governmentwide guidance, including concerning protection of privacy and of sensitive law enforcement, national security, and other protected information, that the AI's entry in the use case inventory provides accessible documentation in plain language of the system's functionality to serve as public notice of the AI to its users and the general public. Where people interact with a service relying on AI and are likely to be impacted by AI, agencies must also provide reasonable and timely notice about the use of the AI and a means to directly access any public documentation about it in the use case inventory.

As shown, the Memorandum M-24-10 imposes a number of obligations to the authorities, which are aimed at ensuring the protection of citizens' rights. As such, for example, the introduction of the right to a human oversight on decisions taken through AI technologies is emblematic of the US government's effort to protect the citizens' rights. This is also considering that some decisions will not be taken using AI systems if additional human oversight, intervention, and accountability cannot be guaranteed. Additionally, the performance of a complete AI impact assessment can result very useful in addressing, before the use of AI in public decisions, the impact of the technology on citizens' rights. In the same direction, ensuring adequate human training and assessment is another important element that contributes to ensure that the public personnel is well trained and familiar with the use of AI technologies, specifically when the their use can have an impact on citizens exercise of civil rights.

Overall, the analysis of the AI Executive Order and of the Memorandum M-24-10 shows some similarities with the obligations and rights that the AI Act provides for the use of high-

risk systems. However, in the case of the US it should be noted that the above analyzed provisions will apply to all the AI systems used by the US agencies, regardless of their specific risk level. Therefore, considering the public sector, the AI Executive Order's approach seems to be more comprehensive than the one used in the EU. However, if not backed up by more concrete legislative measures, the US actual provisions are unlikely to have such a virtuous impact.

2.2 Public Procurement of AI

As the AI Act takes into account the risks related to the chain of value of AI systems, similarly, the AI Executive Order considers the procurement issues in section 2. In this case, the interesting element is the attention posed by the US government in AI risk management specifically in the field of public procurement. Particularly, as anticipated in the previous paragraph 2 of this chapter, the AI Executive Order asks for external (or relational) governance requirements that revolve around the agency's ability to control and challenge tech providers. Moreover, Memorandum M-24-10 also establishes a series of recommendations for managing AI risks in the context of Federal procurement. This involves the transfer (back-to-back) of minimum risk-management practices to AI contractors but also includes commercial considerations. In this specific field, the tone of the AI Executive Order as well as of the Memorandum M-24-10 indicates that this set of requirements is meant to neutralize risks of commercial capture and commercial determination by imposing oversight and external verification.

From an AI procurement governance perspective, the requirements in Section 10.1(b)(viii) are particularly relevant. As some of those requirements will need further development with a view to their operationalization, Section 10.1(d)(ii) of the AI Executive Order requires the Director of OMB to develop an initial means to ensure that agency contracts for the

acquisition of AI systems and services align with its Section 10.1(b) guidance[458]. Accordingly, the guidance required by Section 10.1(b) of the AI Executive Order has been formulated in the Memorandum M-24-10 which offers more detail on the relevant governance mechanisms and the requirements for AI procurement[459]. Indeed, it relies on a tiered approach to AI risk by imposing specific obligations in relation to safety-impacting and rights-impacting AI only. It lays the foundations for a significant strengthening of the governance of AI procurement with a view to embedding safeguards in public sector AI use. A crucially important characteristic in the design of these governance mechanisms is that the Memorandum M-24-10 imposes significant duties on the agencies seeking to procure and use the AI, and it explicitly seeks to address risks of commercial capture and commercial determination. Another crucially important characteristic is that, at least in principle, the use of AI is made conditional on compliance with a rather comprehensive set of preventative and in-use risk mitigation measures.

The general aspects of this governance approach, thus, offer a very valuable blueprint for other jurisdictions considering how to boost AI procurement governance. Particularly:

1. regarding the practices for managing risk and performance for rights-impacting AI and safety-impacting AI, the Memorandum M-24-10 asks for existing contracts for AI systems or services where an agency use is rights-impacting or safety-impacting, agencies to update contract terms as needed to comply with applicable requirements. This includes updates to contractual terms where an AI system or service was initially acquired for a use that did not impact rights or safety, but subsequently is used or planned to be used in a manner that does impact rights or safety;

[458] *Some thoughts on the US' Executive Order on the Safe, Secure, and Trustworthy Development and Use of AI,* How to crack a nut, A blog on EU Economic Law, (November 7th, 2023).
[459] *Ibid*.

2. agencies must cease use of AI systems or services that impact rights or safety in cases where required risk management practices cannot be sufficiently implemented, as determined by the agency. Additionally, agencies should incorporate transparency requirements into contractual terms and solicitations to obtain necessary information and access. Agencies must ensure that vendors provide them with the information and documentation necessary to monitor the performance of an AI system or service and implement applicable requirements. This may include information about the AI's functionality and use that may be publicly posted in the agency's AI use case inventory. Furthermore, careful consideration should be given to the range of potential agency use cases for the acquired AI system or service, and how the information required to facilitate compliance may depend on whether vendors are developers or deployers of an AI system or service;
3. agencies must consider whether any or all of the following categories of information must be provided by the vendor to satisfy the requirements of Memorandum M-24-10 or to meet the agency's objectives: (i) performance metrics, including real-world performance for specific sub-groups and demographic groups to surface discriminatory outcomes; (ii) information about the training data, including the source, provenance, selection, quality, and appropriateness and fitness-for-purpose of the training data, the input features used, time period across which training data was collected, and any filters used; (iii) information about programmatic evaluations of the AI system or service, including the methodology, design, data, and results of how the evaluation of the program delivering the AI system or service was conducted; (iv) information about testing and validation data, including the source, provenance, quality, and appropriateness and fitness-for-

purpose of the testing and validation data, the time period across which it was collected, and the extent of overlap or other possible lack of independence from training data; (v) information about how input data is used, transformed, and retained by the AI and whether such data is accessible to the vendor; (vi) information about the AI model(s) integrated into an AI system or service, including the model's version, capabilities, and mitigations, to the extent it is available to the vendor; (vii) the intended purpose of the AI system or service, known or likely unintended consequences that may occur when deployed for the intended purpose, and known limitations; and (viii) data protection metrics or assurance indicators for data in transit and at rest in AI systems. If the agency must obtain any of this information to satisfy legal or policy requirements, then the agency must incorporate requirements for the submission of that information into solicitation and/or contract documents;

4. agencies must delineate responsibilities for ongoing testing and monitoring and build evaluations into vendor contract performance. Agencies must ensure that contractual terms provide the ability to regularly monitor and evaluate (*e.g.*, on a quarterly or biannual basis, based on the needs of the program) performance and risks throughout the duration of the contract. To do so: (i) agencies must use data defined by the agency (*e.g.*, agency validation and testing datasets) when conducting independent evaluations to ensure the AI system or service is fit for purpose; (ii) contracts must require vendors to provide agencies with sufficient access and time to conduct any required testing in a real-world context, including testing carried out by others on behalf of or under agreement with the agency. Alternatively, agencies may require a vendor to regularly provide the results of an AI system or service's testing in a real-world operational context and the

benchmarks used, with sufficient detail such that the testing could be independently verified or reproduced, if practicable; (iii) contracts must not prohibit agencies from disclosing how they conduct testing and the results of testing; (iv) contracts must detail the examination, testing, and validation procedures the vendor is responsible for and the frequency with which they need to be carried out; (v) where appropriate, agency contracts for AI systems or services must also include terms that require vendors to provide the government with the results of performance testing for algorithmic discrimination, including demographic and bias testing, demographic characteristics of groups the performance testing has been conducted on, or third-party evaluations and assessments providing an equivalent level of detail. Alternatively, agencies may require a vendor to provide the results of performance testing to address these issues; and (vi) agencies must also consider how testing and monitoring, including as part of post-award management, impacts financial planning and budgeting requirements;

5. agencies must set criteria for risk mitigation and prioritize performance improvement. Agencies must have the ability, throughout the entire lifecycle of the contract, to update risk mitigation options and prioritize performance improvement of the AI system or service. To do so, agencies should consider: (i) contractual terms that require vendors to regularly monitor an AI system's performance and rectify any unwanted system behavior, such as retraining the model or adding additional mitigations to the system, based on performance or event-based triggers; (ii) contractual terms that require vendors to meet performance standards before deploying a new version of its AI system or service in performance of an agency contract or for a vendor to roll-back to a previous version if a new version fails to meet performance standards and requirements; (iii) incentivizing improved

model performance through performance-based contracting and incentive contracts; (iv) contractual terms requiring vendors to participate in program evaluations sponsored by the agency to assess implementation and effectiveness, as an additional incentive to assess and improve the intended use case of an AI system, or service; and (v) contractual language that requires vendors to document tools, techniques, coding methods, and testing results, as a means of promoting interoperability and mitigating vendor lock in;

6. agencies must require AI incident reporting. In addition to any existing reporting requirements for cybersecurity or other security-related incidents and breaches of Personally Identifiable Information (PII), agencies must, to the greatest extent practicable, include contractual terms requiring that vendors have a process for identifying and disclosing to agencies serious AI incidents and malfunctions of the acquired AI system, or service within 72 hours, or a timely manner based on the severity of the incident, after the vendor reasonably believes the incident occurred;
7. contracts must include any requirements for additional access, information, or necessary documentation about the AI system or service necessary for the agency to carry out any plans for notice and appeal procedures. Agencies should also consider whether, to ensure successful implementation of a notice and appeal process, the contracts must specify the timeframes in which critical information will be provided by the vendor.

As shown, the attention paid by the US government to contractual relations is evident, especially when the contracting party is the public actor. In this sense, the approach of the US government seems to be similar to the one adopted by the European legislator in the EU model contractual AI clauses to pilot in procurements of AI[460].

[460] Please see the analysis in chapter IV of this work.

2.3 Evaluations of the US approach used in the AI Executive Order

In any case, even though the AI Executive Order and the subsequent Memorandum M-24-10 acknowledge the AI-related concerns and civil rights risks specifically connected to the use of AI by the public actor, as well as specific provision in the field of public procurement, it is necessary to carry out evaluations on the practical efficacy of the analyzed provisions.

Particularly, in assessing the effectiveness of the system, the first element to be considered is that the AI Executive Order frequently relies on soft law terminology, best practice principles, and guidelines. Indeed, the AI Executive Order often uses legally non-binding terminology. As such, it contains ambiguities, grey areas, loopholes, and discretionary leeway. Furthermore, it relies significantly on best practice principles, benchmarks, and guidance, many of which have still to be drafted, remaining unclear how the standardization process will work and how adequate stakeholder engagement could be ensured.

Additionally, closely related to the previous point of criticism is the AI Executive Order's apparent lack of monitoring as there is no separate government agency – *e.g.*, an AI commission or board, as suggested by the AI Act – that could oversee compliance and sanction non-compliance. Consequently, the act misses an adequate and effective governance regime, *i.e.*, monitoring, enforcement, and sanctioning mechanisms, having an impact on its efficacy in providing binding rules. There is also limited enforcement and sanctioning. The AI Executive Order resembles more a soft-law than a hard-law document. It also does not mention administrative fines or penalties – or other behavioral or structural remedies – similar to the ones introduced by the AI Act.

Moreover, it should be considered that it also lacks a risk-based framework, similar to the

one introduced by the AI Act. This aspect, while it may have a positive effect on expanding the application of the provisions, it does not provide any kind of exclusion in relation to the use of AI systems for certain purposes that may have a strong impact on citizens' rights (*e.g.*, constant biometric monitoring or social credit scoring), neither some kind of prioritization of specific aims other than AI rights' related concerns.

In addition, the AI Executive Order also significantly relies on company self-monitoring and industry self-regulation (*e.g.*, in terms of accountability and transparency and certification and standardization). For instance, the AI Executive Order requires developers of the most powerful AI systems [to] share their safety test results and other critical information with the U.S. government. Yet, it remains unclear what this exactly entails.

Furthermore, its level of coherency and stringency could be improved, for instance, by removing redundancies and better connecting or integrating the various guiding principles.

Lastly, it remains elusive about its implementation. Section 12 remarks that a White House AI Council will be established within the Executive Office of the President. The Council's function is 'to coordinate the activities of agencies across the Federal Government to ensure the effective formulation, development, communication, industry engagement related to, and timely implementation of AI-related policies'. Yet, no further details are provided, nor is there any mention of democratic accountability, judicial (and congressional) oversight, and/or possible procedural rights (*i.e.*, remedy and complaint mechanisms)[461].

Therefore, these elements make it not easy to understand the actual implications of this landmark initiative.

[461] *Some thoughts on the US' Executive Order on the Safe, Secure, and Trustworthy Development and Use of AI,* How to crack a nut, A blog on EU Economic Law, (November 7th, 2023).

4. Memorandum on Harnessing Artificial Intelligence to Fulfill National Security Objectives

Additionally, in light of the AI Executive Order, on October 24th, 2024, the Biden Administration issued a new document titled as 'Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence' (Memorandum on AI for national security objectives), which has strengthened the discipline regarding the use of AI, especially when such tools are used by the public actor in the specific field of the national security. Although this work is not focused precisely on the use of AI systems in national security interests (mostly also because the AI Act excludes its application to AI systems used for national security purposes), its brief analysis could result interesting as to understand the additional safeguards the US government has taken in a potential highly intrusive use of AI systems by the public actor.

The purpose of the document, according to the previous US administration, is to galvanize federal government adoption of AI to advance the national security mission, including by ensuring that such adoption reflects democratic values and protects human rights, civil rights, civil rights and privacy. The Memorandum on AI for national security objectives fulfills a directive set forth in the AI Executive Order. Specifically, the directive called for a memorandum to 'address the governance of AI used as a component of a national security system or for military and intelligence purposes', including taking into 'account current efforts to govern the development and use of AI for national security systems'. Therefore, the document reflects these purposes, and it includes numerous references to democratic principles and the threat of digital authoritarianism[462]. More specifically, the Memorandum

[462] S. Feldstein, *The Rise of Digital Repression*, Oxford University Press, 2021. Moreover, to further explanation on this term, please see section 3.1 of chapter II of this work.

seeks to balance national security objectives with human rights, civil rights and civil liberties that could be harmed through the use of AI technologies and systems by the military and intelligence services. Such concerns are not just limited to the national security context, however, and were identified by the administration in the AI Executive Order as a risk for all federal agencies, requiring the OMB to develop guidance for federal agencies to strengthen AI governance, advance responsible AI innovation, and manage risks from the use of AI.

Analyzing the Memorandum on AI for national security objectives, it is possible to affirm that its aim is to not only develop the systems, safeguards, and processes required to harness AI technologies that advance the US national security mission but also to lead the world in the responsible application of AI as it relates to national security. The Memorandum on AI for national security objectives states that AI in the national security context can offer great benefits, but misuse – even without malintent – threatens to bolster authoritarianism worldwide, undermine democratic institutions and processes, facilitate human rights abuses, and weaken the rules-based international order[463].

According to the Memorandum on AI for national security objectives, in order for the US to facilitate a stable and responsible AI governance landscape internationally, it is important that it make use of these systems in a way that protects human rights and civil liberties as well as privacy and safety. But it must do so with responsible speed or risk losing ground to strategic competitors. This requires the US government and its national security institutions to make meaningful changes to their strategies, capabilities, informational infrastructure, governance, and organization so that AI – as it becomes increasingly more general-purpose and affects nearly all domains with national security significance – is not relegated to a single institutional silo. Moreover, the Memorandum on

[463] G. Miller – B. Lennett, *White House AI Memo Promises to Balance National Security Interests with Privacy and Human Rights*, Tech Policy.Press, (October 24th, 2024).

AI for national security objectives sets out three objectives to guide its activities relating to AI and national security:

1. lead the world's development of safe, secure, and trustworthy AI;
2. harness powerful AI with appropriate safeguards;
3. cultivate a stable and responsible framework to advance international AI governance[464].

Much of the stated objectives align with AI Executive Order. Indeed, it reasserts that the government must provide safety and security guidance to AI developers and users and rigorously assess and help mitigate risks that AI systems might pose. It also defines success as measured not only by US technological innovation, but also by its leadership in developing global norms rooted in international law, human rights, civil rights, and democratic values, as well as it warns not to take the unmatched vibrancy and innovativeness of the US AI ecosystem for granted so that it can be proactively strengthened and remain the most attractive destination for global talent and home to the world's most sophisticated computational facilities. Moreover, the US must proactively construct testing infrastructure to assess and mitigate AI risks to realize AI's positive potential and preserve US AI leadership. The Government wants to pursue tools for reliably testing AI models applicability to harmful tasks and deeper partnerships with institutions in industry, academia, and civil society. Commerce, acting through the AI Safety Institute (AISI) within the National Institute of Standards and Technology (NIST), will be the primary point of contact with private-sector AI developers. They will facilitate voluntary pre- and post-public deployment testing for safety, security, and trustworthiness of frontier AI models and lead voluntary unclassified pre-deployment safety testing, including risk assessments related to cybersecurity, biosecurity, chemical weapons, and

[464] *Ibid*.

system autonomy.

Lastly, the Memorandum on AI for national security objectives makes clear that AISI's direct responsibilities do not extend to the evaluation of AI systems for the potential use by the United States Government for national security purposes, as these responsibilities lie within the agencies considering such uses[465].

Overall, this Memorandum represents one of the first attempts by a democratic country to design a specific framework in the use of AI for such valuable interests, such as the ones pursued for national security issues. However, the document relies on soft law terminology, best practice principles, and guidelines, which in the end have just a programmatical effect.

5. Comparing US and EU approaches in regulating AI

Before highlighting the main differences between the Executive Order and the AI Act, it is worth noting the similarities that exist between them as well.

Particularly, the most interesting intersection definitely concerns privacy and data protection, which also appears as the most challenging aspect as far as the US is concerned. The AI Executive Order focuses on privacy both in the principles listed in Section 2 and in the specific Section 9. On one hand, it acknowledges that AI 'is making it easier to extract, re-identify, link, infer, and act on sensitive information about people's identities, locations, habits, and desires', leading to increasing risk of exploitation and exposure of personal data. Therefore, the AI Executive Order makes clear the commitment of the Federal Government to ensure that 'the collection, use, and retention of data is lawful, is secure, and mitigates privacy and confidentiality risks'. Interestingly enough, the AI Executive Order encourages resort to privacy-enhancing technologies

[465] *Ibid*.

among the available technical tools. This way, it makes visible its reliance on the power of technology as a modality of regulation. On the other hand, Section 9 further elaborates in greater detail this commitment, reiterating the acknowledgment that AI can facilitate the collection or use of personal information or the making of inferences about individuals. It is difficult not to see, along these lines, an implicit but warm urge to Congress to pass data protection legislation. As noted by some privacy scholars[466], this would significantly help put into practice the mandates established under the Executive Orders (including mandates other than that under Section 9). However, it seems unlikely that an ambitious goal such as setting comprehensive data privacy legislation can be achieved at this historical moment. Also, the California Consumer Privacy Act has planted the seeds for the flourishing of state statutes aimed at empowering citizens against the processing of their data most notably by businesses: because of this wave, it is unlikely that Congress will take the floor on a matter on which it was traditionally reluctant to step foot. This is perhaps a key sector where EU law could actually play an influence and produce its Brussels effect: not the AI Act, but rather the General Data Protection Regulation, which already applies to US-based organizations to the extent they process data of European residents when offering them products or services or monitoring their commercial behavior in the EU[467].

Analyzing the differences, generally, it is undoubtable that the EU and the US have used a different approach in regulating AI. The lack in the US legal order of a comprehensive federal law specifically governing AI is probably the first and clearest difference as, in the EU, the AI Act has been recently enacted[468]. Indeed, while the AI Act is legally binding,

[466] G. Zanfir-Fortuna*, Four Data Protection Threads in Today's Biden-Harris EO on AI, through a Global lens*, (October 30th, 2023).
[467] M. Bassini, *The Global Race to Regulate AI: Biden's Executive Order Spillover Effects on the EU AI Act*, Medialaws, 2023.
[468] *Ibid*.

comprehensive framework regulating private and public sectors across the EU, in contrast, the AI Executive Order is non-legally binding and primarily focuses on centralizing AI efforts within the federal government. An Executive Order is thus different from other (legally binding) hard laws, such as the AI Act, which includes establishing government bodies responsible for compliance monitoring and introducing penalties and fines for non-compliance (*i.e.*, behavioral or structural remedies)[469].

Besides that, it should be noted that the (previous) US approach to AI regulation relies on two primary strategies: establishing guidelines and standards through federal agencies and encouraging self-regulation within the industry. In this direction, the US approach is guided by several overarching objectives, including fostering openness and competitiveness in the AI-driven economy, enhancing safety while managing risks, and maintaining a technological edge over global competitors. These strategies and objectives have shaped the US approach to AI regulation, which is characterized as market-driven, sector-specific, and primarily vertically[470] focused. It is highly decentralized across federal agencies and grounded in a risk management framework. Unlike the state-driven model, which exercises government control over AI development to direct economic policies and maintain political stability, and the EU's rights-driven framework, which focuses on safeguarding fundamental rights through comprehensive regulations like the AI Act, the US embraces a market-driven approach that minimizes state intervention, promoting innovation and economic growth through voluntary standards and self-regulation. The belief is that excessive regulation could hinder technological advancement. As such, the government and private organizations have published their AI ethical principles and

[469] D. Tatevik, *The U.S. Approach to AI Regulation: Federal Laws, Policies, and Strategies Explained*, (September 09th, 2024).

[470] Please note that, in a vertical strategy, policymakers take a bespoke approach, creating different regulations to target different applications or types of AI (US approach). In a horizontal approach, regulators create one comprehensive regulation that covers the many impacts AI can have (EU approach).

guidelines, encouraging the industry to adopt responsible practices while maintaining flexibility in the rapidly evolving technological landscape[471]. Following this approach, the characteristic of US AI regulation is its decentralization, allowing different federal agencies to tailor regulations to specific sectors and use cases. Therefore, different from the EU's centralized and 'horizontal', strategy, exemplified by the AI Act, which sets broad standards applicable across various AI applications and industries, the US approach is more fragmented. Indeed, the AI Act applies uniformly across all AI systems and sectors, establishing a comprehensive, cross-sector regulatory framework that ensures consistency across EU Member States. This unified approach aims to harmonize standards and practices, requiring all AI applications to adhere to the same comprehensive rules although utilizing a multitier risk-based approach. In contrast, the US adopts a decentralized, sector-specific strategy where different federal agencies have jurisdiction over specific aspects of AI based on their existing mandates[472,473].

Moreover, the AI Executive Order sets different priorities than the AI Act. It focuses, for instance, much more on safety and security aspects in connection with innovation and competition. It also covers a variety of topics that are currently neglected in the AI Act, namely worker and consumer protection, competitiveness, innovation especially in the

[471] Notably, in July 2023, the Biden-Harris Administration secured voluntary commitments from major AI companies, including Amazon, Google, Meta, and Microsoft, to enhance AI safety, security, and transparency. These companies agreed to rigorous testing, sharing safety protocols, reporting vulnerabilities, and developing tools to identify AI-generated content. They will also address societal impacts like bias and privacy and explore AI's potential to tackle global challenges such as climate change and healthcare.

[472] Noteworthy is that it is possible to observe these different regulatory traditions or pathways in other economic or policy areas. For instance, the US still lacks a federal privacy or data protection law. Instead, it relies primarily on sector or industry-specific regulations (such as HIPAA), FTC's guidelines (such as 'Notice and Choice'), and state laws (such as the CCPA/CPRA). The focus is primarily on voluntary or industry self-regulation, soft laws, and best practice guidelines instead of mandatory or legally binding regulation. The EU, however, follows a more comprehensive and holistic approach with the GDPR. It relies on mandatory laws and regulations (*i.e.*, hard laws) and considers privacy and data protection as fundamental rights (*i.e.*, E.U. citizens have the right to access their data, correct inaccurate data, and demand the deletion of specific data, according to the so-called 'right to be forgotten' or 'right to erasure').

[473] M. Bassini, *The Global Race to Regulate AI: Biden's Executive Order Spillover Effects on the EU AI Act*, Medialaws, 2023.

form of supporting SMEs, and leadership in AI research and development. While some of those issues are covered in other European laws and regulations (*e.g.*, the EU's Digital Markets Act deals with competition and antitrust issues), it would have been worth revisiting those aspects and discussing them in more detail in the AI Act (for instance, the AI Act touches upon promoting AI research and development and innovation and supporting SMEs in the sections devoted to so-called 'regulatory sandboxes' but not elsewhere). The AI Act could thus learn from the AI Executive Order by focusing more on consumers, workers, and SMEs entrepreneurs[474].

Furthermore, while the AI Act directly applies to all Member States, impacting the private sector directly, the US seems to rely on executive orders to 'bypass' the need for comprehensive legislative action to regulate private industry directly. These orders instruct executive departments to develop industry-specific standards and guidelines for AI development and governmental use.

Lastly, both the US and EU align conceptually to sharing principles of trustworthy AI and endorsing international standards. However, they use a different approach in the application of these principles. Indeed, in the AI Act higher-risk AI applications that have the potential to cause significant harm are subject to more stringent regulations, while less stringent requirements govern lower-risk applications. As such, the specifics of their AI governance regimes vary significantly, with the EU's approach being more coordinated and binding. Indeed, the EU's AI Act employs a more rigorous and targeted framework with a tiered risk classification system categorizing AI systems into four levels – minimal, limited, high, and unacceptable – while the US approach generally balances the potential benefits of AI with its risks. As a result, higher-risk applications are likely to undergo greater scrutiny. Thus, the US favors a decentralized, mainly vertical, and market-driven approach

[474] M. Wörsdörfer, *Biden's Executive Order on AI and the E.U.'s AI Act: A Comparative Computer-Ethical Analysis,* Philosophy & technology, 2024, p. 74.

that encourages innovation and is adaptable to the specific needs of different sectors. It is argued that the US AI policy has been - and is likely to remain – 'a mosaic of individual agency approaches and narrow legislation rather than a centralized strategy'[475,476]. In this context, despite the lack of directly prescriptive rules and its limited enforceability as it is, the AI Executive Order draws sufficiently broad guidelines and perhaps is likely to have a more comprehensive and huge impact than the AI Act from a social and economic perspective. The AI Act primarily aims to categorize AI systems into different classes of risk, which reflects the essence of the risk-based approach. It adheres to a regulatory approach that focuses on the use of technology rather than on technology *per se*. This methodology, however, may also lead to undesirable consequences given its lack of flexibility combined with the unprecedented development of AI systems. Instead, the AI Executive Order takes a different approach, which reflects its legal nature. Aimed to promote standards and best practices, including in the context of regulated sectors, and based on extensive stakeholders' consultation, it seeks to accommodate the desire of industry (but also of the federal government) to unleash the full potential of AI. Consistently with the US understanding of the role of innovation at the intersection with fundamental rights, this goal is not achieved – as opposed to the EU – by laying down a comprehensive and detailed set of obligations, but rather by setting binding guidelines and only in a residual way resorting to regulation (for instance, with respect to generative AI). This will not prevent, in any case, the competent authorities for the relevant sectors that have been empowered by the AI Executive Order from setting more prescriptive and detailed rules and enforcing them, similarly to what could happen in the EU, where the design of proper governance for AI is still debated. However, what can be predicted is that most likely in

[475] H. Pouget – M. O'Shaughnessy, *Reconciling the U.S. Approach to AI*, Carnegie Endowment for International Pease, (May 3rd, 2023).
[476] D. Tatevik, *The U.S. Approach to AI Regulation: Federal Laws, Policies, and Strategies Explained*, (September 09th, 2024).

the US a 'holistic view' on AI regulation will be (at least) partially missing, given the number of authorities called to action by the AI Executive Order. The US approach marks in any case a significant departure from the 'regulatory anxiety' of EU lawmakers *vis-à-vis* disruptive technology and follows the general skepticism in the US legal culture about the role of regulation, most notably when it comes to emerging technologies and possible interferences with human rights. Moreover, the scope of application of the AI Executive Order overall extends far beyond the purely AI domain. While the EU mostly focuses on the establishment and functioning of the internal market and on the implications of AI systems for data protection, the AI Executive Order takes the regulation of AI as an expedient for an American 'gold-plating' that, if realized, will lead to substantial reforms in the social and productive sphere. In EU law, gold-plating describes the controversial practice of Member States to extend the scope of EU directives (which set obligations for Member States to fulfill certain objectives but leave them room to determine the appropriate measures) when transposing them into domestic legal systems. In the context of the AI Executive Order, it is likely that US governmental agencies and – why not? – Congress will do something similar, taking the regulation of AI as the occasion to effect profound reforms mirroring the variety of economic and social sectors impacted by the AI revolution. The cross-sector approach adopted by the AI Executive Order parallels in fact the many diverse industries and areas where AI is supposed to have an impact. Also, despite the AI Executive Order's aim to foster American leadership in the race to AI, it does not provide any 'Washington effect'[477]: unlike the AI Act, the AI Executive Order does not *per se* target non-American entities and its focus lies entirely on American companies and the federal government. However, it calls for a strong international framework to govern the development and the use of AI. It goes without saying that also the upcoming

[477] L. Floridi, *On the Brussels-Washington Consensus About the Legal Defnition of Artifcial Intelligence*, Philosophy & Technology, 2023, p. 1–9.

developments in the EU legal framework, and most notably the possible changes regarding the legal status of foundation models, will have huge impact on US-based companies. Therefore, the interconnections between the two legal orders, as well as the different approaches adopted are stronger than ever. These factors, therefore, called for a necessary comparative analysis of the two acts[478].

5.1 The AI Executive Order and the AI Act in regulating the use of AI by the public actor

Considering the different approaches in regulating AI by the US and EU, comparisons need to be drawn also in relation to the different regulations established in relation to the use of AI technologies by the public actor. Both frameworks aim to promote responsible AI innovation, safeguard democracy, and protect human rights, with a focus on fairness and transparency. However, the approach used differs clearly.

Particularly regarding the UE legal order, it should be taken into account that the AI Act does not have a specific section that relates to the use of AI by the public actor, as the regulation applies a risk-based approach to providers, deployers, importers, distributors and users independently they are public or private entities. That being said, considering that, depending on the circumstances regarding its specific application, use, and level of technological development, AI may generate risks and cause harm to public interests and fundamental rights that are protected by Union law, the AI Act classifies some AI systems that are used for public functioning as high-risk systems. Therefore, in specific cases there are obligations that have to be fulfilled by the public actor[479].

On the other hand, the AI Executive Order and the Memorandum M-24-10 provide for

[478] M. Bassini, *The Global Race to Regulate AI: Biden's Executive Order Spillover Effects on the EU AI Act*, Medialaws, 2023.

[479] To further explanation on this theme, please see section 3.3.5 and 3.3.6 of chapter III of this work.

extensive rules on the use of AI by the public actor (*i.e.*, federal agencies), giving comprehensive guidelines both in the areas of the protection of citizens' rights and in the public procurement one, however without imposing binding provisions. Therefore, while the AI Act seems to provide an enforceable regulatory framework without focusing on the public sphere, the AI Executive Order offers a flexible, government-focused approach, allowing more straightforward adaptation to technological advancements, which however has not any binding effect rather than a programmatic one[480].

6. Concluding Remarks

As shown, this chapter has conducted a legal analysis of two of the world's premier AI legislations – the AI Act and the AI Executive Order. It has identified several strengths (*e.g.*, the transition from soft to semi-hard law) and weaknesses (*e.g.*, ineffective governance) of both initiatives, as well as parallels (*i.e.*, similar points of criticism) and differences (*i.e.*, different regulatory traditions and approaches). Besides, the chapter has acknowledged the need to strengthen both initiatives, such as hardening enforcement, monitoring, and sanctioning.

In the highlighted scenario, it will also be interesting to see whether another 'Brussels Effect' will be created, as it has already happened with the GDPR (and might happen with the Digital Services Act and Digital Markets Act) or whether the 'Washington Effect'[481] will prevail (*i.e.*, countries following the US instead of the EU model and focusing more on semihard law)[482].

[480] D. Tatevik, *The U.S. Approach to AI Regulation: Federal Laws, Policies, and Strategies Explained*, (September 09th, 2024).
[481] L. Floridi, *On the Brussels-Washington Consensus About the Legal Defnition of Artifcial Intelligence*, Philosophy & Technology, 2023, p. 1–9.
[482] M. Wörsdörfer, *Biden's Executive Order on AI and the E.U.'s AI Act: A Comparative Computer-Ethical Analysis,* Philosophy & technology, 2024, p. 74.

CHAPTER VI: THE FINDINGS OF THE RESEARCH

Contents: 1. The investigated questions; 2. The AI-driven public actor from a fundamental rights perspective: the findings of the research; 3. The interconnections between the public and the private actor: the findings of the research; 4. The US legal order's answers: the findings of the research; 5. Final remarks.

1. The investigated questions

As described in this research, in the last few years, with the increasing use of technology, and especially of AI, new aspects in the exercise of the public actor's power can be observed. Specifically, as AI technologies begin to play a dominant role in the contemporary exercise of power, it becomes increasingly important to examine the phenomenology of a new kind of power and its unique challenges to fundamental rights[483]. Therefore, as analyzed, the so-called AI-driven public actor has elements of absolute novelty that have been assessed in consideration of the UE legal order – and briefly of the US one –, also thanks to the introduction of the AI Act in the EU, as well as of the AI Executive Order in the US.

In this context, the previous chapters have tried to answer two major questions. First, which is the impact of the new paradigm of the AI-driven public actor on the exercise of the citizens' fundamental rights? This core question requires answering the following sub-question: which are the limits and the guarantees that the European Union (EU) legal order is currently ensuring within this new paradigm? Second, does the public actor's procurement of private AI gives to the private actor a specific role in public decisions? If yes, has the EU legal order established any provisions to balance the role of the private actor in public decisions? Moreover, this work – albeit briefly – researches whether the US

[483] A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33.

legal order has enacted any provision dealing with these issues.

Overall, the critical constitutional aspects of the algorithmisation of the public actor have been examined, as well as whether a deterioration in the relationship between the state and the citizen can be witnessed, as individuals are increasingly subjected to fully automated decisions that have a strong impact on the exercise of their fundamental rights. Accordingly, the answers to these questions have important implications in terms of understanding the future of the application of fundamental rights. Indeed, the future of the guarantee of fundamental rights, and perhaps the future of democracy, depends in large part on the ways in which new technologies are changing the relationship between the public actor, the private one and the civil society[484]. On this matter, while recognizing the good work already being carried out in the AI law (as evident in the literature identified in this research), this consolidated analysis of issues hopes to further provide insights and add to the much-required need for further and sustained discussions on this topic, given AI's increasingly widespread deployment and use and the gravity of its impacts on individuals and their fundamental rights[485], specifically in the public field.

2. The AI-driven public actor from a fundamental rights perspective: the findings of the research

As the first question regards the investigation of the impact of the new paradigm of the AI-driven public actor on the exercise of the citizens' fundamental rights and which are the protection that the EU legal order is currently trying to guarantee, this work focused primarily on the European legislation.

[484] D. Tiberiu – Y. Lupu, *Digital authoritarianism and the future of human rights*, International Organization, 2021, p. 991-1017.
[485] R. Rodrigues, *Legal and human rights issues of AI: Gaps, challenges and vulnerabilities*, Journal of Responsible Technology, 2020.

As such, this research in chapter III has shown how the limitations of the current incarnation of the right to good administration, as well as of the other fundamental rights[486] in the CFREU are not sufficiently addressed through the adoption of secondary legislation (*i.e.*, GDPR, AI Act), nor through case law.

Indeed, the use of AI systems within the AI-driven public actor still raises concerns from a fundamental rights perspective, as from a constitutional and legislative view it entails a possible detriment to citizens' rights. This outcome emerges from the 3-layered analysis that has been carried out in chapter III, focusing on the constitutional rights and legislative framework of AI, as well as on the answers provided by the administrative judges. On this merit, chapter III has analyzed the balance between certain (and selected) fundamental rights and the interest of the public actor in using technologies to make the functioning of the public apparatus more efficient. It is without a doubt that fundamental rights law provides an invaluable organizing framework of concrete rights and respective obligations for assessing the use of AI systems in the AI-driven public actor paradigm. Following this path, this work highlighted the risk of: (i) violation of the privacy right and of the data protection one; (ii) discrimination when prediction analytics come into play; (iii) the allocation of the burden of proof on citizens when the state refuses to disclose information about a given AI system in relation to the right to have an effective remedy; (iv) impossibility to guarantee the right to transparency and access to file; (v) impossibility to guarantee a reasoned decision[487]. Additionally, chapter III considered the responses of the European legislator and of national judges. In relation to the solutions provided by the European legislator, it has been found that the GDPR poses several interpretative problems with respect to its applicability to the use of AI systems. Moreover, it provides

[486] Please note that chapter III focused on the right to privacy, to personal data protection, non-discrimination, the right to a good administration and the right to an effective remedy.
[487] A. Rachovitsa - N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.

for a right to explainability that is too limited in relation to the impact of the legislation on citizens' fundamental rights. In light of this, the analysis then focused on the responses of national jurisprudence. In the absence of specific legislative provisions (before the entry into force of the AI Act), it is the administrative judge who has so far played a central role in the topic of public actor's algorithmic decisions. However, as shown, the path toward building a 'technological due process' within the paradigm of the AI-driven public actor still appears to be on a preliminary stage. Finally, the analysis focused on the responses given by the legislator in the AI Act, in which not all the open issues about the protection of fundamental rights within the AI-driven public actor are solved even if the Regulation tries to adopt a 'hybrid' legislative approach according to which the technology has to be fundamental right oriented since its design[488]. Indeed, the emerging European approach to reshaping good administration for digital public governance is significantly constrained due to the threshold issues that result from having placed the regulatory focus on high-risk AI uses, as well as by the absence of strictly enforceable individual rights. This will do little to address the broader gaps in the regulation of good administration. Such gaps are particularly visible when the focus is put on the mass effects that the digitalization of the public sector decision-making is bound to generate. Such mass effects risk depriving good administration rights from any practical effect where supported or automated decision-making systems present administrative decisions as a fait accompli and where the sheer numbers of potentially affected citizens and related claims are bound to overwhelm existing mechanisms for the review or appeal of those decisions[489].

Accordingly, following this path – making the current provisions more oriented to the hybrid legislative approach – is actually the only possible way is to ensure that the value (*i.e.*, the

[488] A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33.
[489] A. Sanchez-Graells, *Resh (AI) ping Good Administration: Addressing the mass effects of public sector digitalization*, Laws, 2024, p. 9.

meaning) of protecting the fundamental rights of the citizens as their interests become an integral part of the development of AI. Consequently, this calls for an advanced legal framework that not only governs the use of digital and automated systems but also ensures their congruence with the ethical and cultural fabric of each unique jurisdiction. This approach advocates for the development of regulatory mechanisms that are both context-sensitive and responsive to the dynamic nature of digital technologies, ensuring that they serve the public interest and enhance democratic values. In addition, a holistic perspective of public law should ensure that fundamental rights and public values are considered more closely in national digitalization strategies. Public values represent the foundational ideals, ethical standards, and principles that constitute the bedrock of a society's collective conscience. These values act as benchmarks for action, guiding the public actor in the exercise of their functions and pursuit of the public interest. Public values shape not only the mindset of civil servants but also the objectives towards which governmental institutions should strive[490].

Overall, a final reflection is that such interventions are urgent. While the legislative framework adapts at a slow pace, the public sector is quickly accumulating a stock of data and digital technology supported or automated decision-making solutions that will be very difficult to dismantle once it gets embedded. More importantly, the accelerating process of digital transformation is currently externalizing significant risks on citizens and, most likely, disproportionately on the most vulnerable citizens. This mere fact is in itself a threat to the existing obligations to protect and promote a broad array of human and fundamental rights. Much like the EU has been keen to be a trendsetter in the regulation of some uses of AI, it should also be willing to be a trendsetter in shaping good administration for the AI-

[490] S. Ranchordas, *The Invisible Citizen in the Digital State: Administrative Law Meets Digital Constitutionalism*, in (eds.) J. De Poorter – C. Oirsouw – G. Van Der Schyff, *European Yearbook of Constitutional Law*, Tilburg Law School Research Paper, (December 24th, 2023).

drive public actor paradigm[491].

3. The interconnections between the public and the private actor: the findings of the research

In order to have a comprehensive view of the AI-driven public actor paradigm, the research has analyzed the fundamental role that the private actors are having in this transition.

Two aspects of this strong interconnection between public and private actors have been observed. First, more frequently private actors hold the know-how for the creation and use of sophisticated technologies such as AI. Therefore, the public actor has to rely on the provision of technologies by the private sector which, usually, is not open to sharing proprietaries information about the functioning of the AI systems. Second, in order to enhance transparency between the public and the private actor, as well as in order to create positive externalities, the EU is fostering a strategy on data sharing between them, specifically with the aim of fostering trusted mechanisms and services for the re-use, sharing and pooling of high quality.

Analyzing the first aspect, this research has found that the EU's regulatory strategy in relation to public sector AI procurement and deployment has many gaps. Indeed, significant amounts of AI procurements and use cases will remain unregulated even under the AI Act. Moreover, even those high-risk use cases covered by the Act will be subjected to mechanisms of regulation by contract, it can hardly prove effective given the structural limitations in using public procurement as a regulatory tool. The implementation of this model of AI regulation by contract is significantly challenged by the lack of independence of the procurement function and the risks of bypassing of procurement-related constraints by the public sector entities using the AI, their tech providers, or both. It is further

[491] A. Sanchez-Graells, *Resh (AI) ping Good Administration: Addressing the mass effects of public sector digitalization*, Laws, 2024, p. 9.

undermined by the growing gap in the digital skills of many public sector institutions and public buyers. As a result, the risks identified in chapter IV remain largely unaddressed and it would not be sensible to consider that the regulatory adaptation resulting from the AI Act and the pilot model contractual clauses for the procurement of AI will be sufficient, or effective. Particularly, the EU should ensure that no digital technologies are procured or deployed in a way that infringes EU or domestic data and digital regulation, and that no contracting authority procures digital technologies without a sufficient impact assessment and without ensuring adequate digital skills to manage the process and the use of the technology throughout its lifecycle.

As pointed out in chapter IV, the best policy intervention to achieve these goals would involve a mix of: (i) review of the EU procurement rules to create new mechanisms of assurance and, in particular, new mechanisms of impact assessment prior to the launch of a procedure for the procurement of digital technologies, which could be modelled on the impact assessments under the AI Act; (ii) creation of a single handbook of data and digital law applicable to procurement, on open access, curated and permanently updated by the European Commission; (iii) mandatory specific oversight mechanisms under independent authority to be developed at Member State level, to ensure compliance with the impact assessment and other regulatory requirements prior to the launch of the relevant procurement; (iv) coordination of the network of independent national authorities by the European Commission, perhaps through an expansion of the remit and a review of the institutional design of the future EU AI Office; (v) mandatory specific tasks of digital professionalization and recruitment of digital skills into the procurement workforce, or a specialized body tasked with the procurement of digital technologies and a credible plan for its continuous implementation and funding to be implemented by all Member States. The implementation of such a policy mix would be more prescriptive than the strategy

followed to date, whereby the facilitative approach taken by the European Commission in some areas (*e.g.*, procurement professionalization) has yielded limited results. The combined implementation of these measures would significantly minimize the risks and harms arising from inadequate processes of deployment of digital technologies by the public sector, as well as creating a level of protection of fundamental and individual rights, and of collective and social interest, that is homogeneous across the EU[492].

In relation to the second aspect, the relationship between the public and private actors in using AI systems thanks to the creation of a mechanism of data sharing has been analyzed. In this direction, chapter IV has shown that EU data legislation undoubtedly has horizontal – or universal – ambitions. This is visible with the DGA, which attempts to incentivize data sharing in a wide range of scenarios and in particular both market and non-market ones. The horizontality of EU data legislation is also visible with B2G mandatory data sharing that applies irrespective of the type of data, type of public authority and even, to a large extent, irrespective of the type of exceptional need at stake. Moreover, the ambition of EU data legislation to provide for a horizontal regulation of data is visible from the harmonized legal conditions for mandatory data sharing between businesses which applies irrespective of the sector, context and, even more striking, even rationale for mandatory data sharing. Even more important, a set of common features and principles can be identified.

Overall, the EU data sharing legislation is based on the same premise that, as non-rival resources, data shall, in principle, be shared and broadly used in order to optimize their economic potential particularly in the field of AI. Then, the whole of EU data legislation consists of balancing this principle with contextual specificities requiring more closedness. The legitimate rights and interests of different actors – and especially public actor and

[492] A. Sanchez-Graells, *Public Procurement of Artificial Intelligence: Recent Developments and Remaining Challenges in EU Law,* Legal Tech Journal, 2024.

private ones – are symptomatically put increasingly on equal footing. On the one hand, they may tip the scales towards closedness, and, on the other hand, they are interpreted, especially under the DGA, as legal mechanisms likely to support data sharing akin to property entitlements.

However, as pointed out in chapter IV, this framework (and more specifically the DGA) seems incompatible with the many of the provisions of the GDPR[493]. Therefore, despite the fact that the regulatory framework analyzed aims at increasing the exchange of information between the public and private sectors in order to generate positive externalities, implementation difficulties (due to incompatibility with the GDPR) could lead to a clear ineffectiveness of the provisions and, ultimately, to the non-achievement of the objectives declared by the EU.

4. The US legal order's answers: the findings of the research

Lastly, chapter V analyzed the AI Executive Order, even if revoked, also to understand whether the approach used can be considered more effective than the one adopted in the EU.

As previously analyzed, the AI Executive Order and its subsequent act, the Memorandum M-24-10, impose a number of obligations to the authorities, which are aimed at ensuring the protection of citizens' rights. As such, for example, the introduction of the right to a human oversight on decisions taken through AI technologies is emblematic of the US government's effort to protect the citizens' rights. This is also considering that some decisions will not be taken using AI systems if additional human oversight, intervention, and accountability cannot be guaranteed. Additionally, the performance of a complete AI

[493] C. Ducuing, *The Regulation of Data in the European Union: the Data Governance Act and the Data Act*, in (eds.) E. E. Akin – S. Klimbacher – G. Ziccardi, *Smart cities, artificial intelligence and digital transformation law*, Milano University Press, 2024.

impact assessment can result very useful in addressing, before the use of AI in public decisions, the impact of the technology on citizens' rights. In the same direction, ensuring adequate human training and assessment is another important element that contributes to ensure that the public personnel is well trained and familiar with the use of AI technologies, specifically when the their use can have an impact on citizens exercise of civil rights.

Therefore, the analysis of the AI Executive Order and of the Memorandum M-24-10 has shown some similarities with the obligations and rights that the AI Act provides for the use of high-risk systems. However, in the case of the US it should be noted that the analyzed provisions will apply to all the AI systems used by the US agencies, regardless of their specific risk level. Therefore, considering the public sector, the AI Executive Order's approach seems to be more comprehensive than the one used in the EU. However, if not backed up by more concrete legislative measures, the US actual provisions are unlikely to have such a virtuous impact.

Moreover, the attention paid by the US government to contractual relationships is evident, especially when the contracting party is the public actor. Indeed, the AI Executive Order and the Memorandum M-24-10 provide for extensive rules on the use of AI by the public actor (*i.e.*, federal agencies), giving comprehensive guidelines both in the areas of the protection of citizens' rights and in the public procurement one, however, without imposing binding provisions. Overall, chapter V assessed that while the AI Act seems to provide an enforceable regulatory framework without focusing on the public sphere, the AI Executive Order offers a flexible, government-focused approach, allowing more straightforward adaptation to technological advancements, which however has not any binding effect rather than a programmatic one[494].

[494] D. Tatevik, *The U.S. Approach to AI Regulation: Federal Laws, Policies, and Strategies Explained*, (September 09th, 2024).

In the end, in light of the brief analysis carried out on the two AI regulations, a common path can be observed between the US and the EU, despite differences in legal background, scope of application, and oversight systems. Indeed, the content of the two regulations is similar. They both stress the importance of protecting values such as health, safety, democracy, and the rule of law, while recognizing the need to foster and support technological innovation[495].

5. Final remarks

Given the foreseeable pervasiveness of AI used by the public actor, ultimately, this research sought to understand how this new technology must be shaped to support the maintenance and strengthening of fundamental rights, democracy and the rule of rather than weakening it. Indeed, these principles are supreme law of the land – all actions of government, legislators and indeed societal reality are measured against them. As such, the need for framing the future relationship between technology and democracy cannot be underestimated[496].

Despite the inconsistency and the highlighted application issues to date, it is possible to conclude that within the new paradigm of the AI-driven public actor – and more generally, being aware of the disruptive effects of the use of the technologies –, the EU is committed to promoting the ethical and responsible use of AI through regulations and investments that affect the digital environment. To this end, the analysis of AI regulatory instruments suggests an approach aimed at promoting responsible and equitable use of technology. The objective is to create an informed and prepared society for AI. The EU, in particular, aims to establish a reliable and responsible regulatory framework for AI that enhances

[495] C. Scarpellino, *EU and US regulatory approach to AI: a comparative perspective*, LUISS Policy Observatory, 2024.
[496] P. F. Nemitz, *Constitutional Democracy and Technology in the age of Artificial Intelligence (August 18, 2018)*, Royal Society Philosophical Transactions, 2018.

people's lives while preserving societal values. The recently enacted AI Act is a significant step towards addressing the need to limit the potential abuse of AI, even from the public actor. It also acknowledges the importance of striking a balance between regulation and innovation, which is critical to ensuring the responsible and beneficial application of AI.

Towards this direction, there is clearly the intention to introduce a system in which AI should be a human-centric technology and should serve as a tool for people, with the ultimate aim of increasing human well-being. Furthermore, the AI Act specifies that its aim is to enhance the operation of the internal market and encourage the adoption of human-centric and trustworthy artificial intelligence. This must be achieved while maintaining a high level of protection for health, safety, fundamental rights, non-discrimination, against the detrimental effects of artificial intelligence systems in the EU. In this context, to build greater trust in the positive impact that AI can have on society, it is essential that AI and its regulatory framework are developed in accordance with an anthropocentric vision, and therefore with the Union's values as enshrined in Art. 2 of the TEU, the fundamental rights and freedoms enshrined in the Treaties, and the Charter, as stated in Art. 6 TEU.

However, in this context, the impact of the timing of the implementation of the provisions of the AI Act cannot be underestimated. Indeed, as explained in the previous chapters, many of the obligations will apply from 2026. Among others, those relating to high-risk AI systems will have to be applied from 2027. Particularly, given that we may find ourselves in a changing technological environment in the next two years, it cannot be ruled out that the provisions currently enacted will not only prove ineffective and incapable of guiding the path of AI development but will also fail to meet the first of the EU's objectives, *i.e.*, the protection of citizens' fundamental rights.

Therefore, it is important to approach the implementation of the AI Act's provisions with caution and involve a variety of stakeholders to ensure a balance between innovation and

societal principles. By adopting this approach, a harmonious balance between technological advancement and the protection of human well-being can be established. While AI presents exciting technological opportunities, it is important to remember that humans should be the focus of all regulations and their subsequent application. It is crucial to acknowledge that humans are not only the recipients of AI progress but also the architects of it. To ensure that AI contributes to the well-being of society, a balance needs to be struck that requires human intelligence, shrewdness, and logic[497].

[497] A. Pirozzoli, *The Human-centric Perspective in the Regulation of Artificial Intelligence*, European Papers, 2024, p. 105-116.

BIBLIOGRAPHY

- Agenzia per l'Italia Digitale, *Strategia Italiana per l'intelligenza Artificiale 2024-2026*, (July 22nd, 2024).
- AlgorithmWatch, *Automating Society Report 2020*, (October 2020).
- S. Alon-Barkat – M. Busuioc, *Human-AI Interactions in Public Sector Decision-Making: "Automation Bias" and "Selective Adherence" to Algorithmic Advice*, Journal of Public Administration Research and Theory, 2022.
- J. M. Balkin, *Free Speech in the Algorithmic Society: Big Data, Private Governance, and New School Speech Regulation*, University of California, Davis Law Review, 2017, p. 51.
- J. M. Balkin, *The Three Laws of Robotics in the Age of Big Data*, Ohio State Law Journal, 2017.
- S. Barocas – A. D. Selbst, *Big Data's Disparate Impact*, California Law Review, 2016, p. 671.
- M. Bassini, *The Global Race to Regulate AI: Biden's Executive Order Spillover Effects on the EU AI Act*, Medialaws, 2023.
- S. Bertaina – I. Biganzoli – R. Desiante – D. Fontanella – N. Inverardi – I. G. Penco – A. Cosentini, *Fundamental Rights and Artificial Intelligence Impact Assessment: A New Quantitative Methodology in the Upcoming Era of Ai Act*, 2024.
- M. Bertolini – D. Mezzogori – M. Neroni – F. Zammori, *Machine Learning for Industrial Applications: A Comprehensive Literature Review*, Expert Systems with Applications, 2021, p. 114-30.
- A. Bibal – M. Lognoul – A. De Streel – B. Frénay, *Legal Requirements on Explainability in Machine Learning*, Artificial Intelligence and Law, 2021, p. 154–155.
- P. Bizzini, *The algorithm that blew up Italy's school system,* Algorithm Watch, (April 17th, 2023).
- W. Blackstone, *Commentaries on the Laws of England*, 1765-1769.
- N. Bobbio, *Futuro della democrazia*, 1984.
- N. Bobbio, *Liberalism and Democracy*, 1990.
- M. Bogden – A. Rieke, H*elp Wanted: An Examination of Hiring Algorithms, Equity, and Bias*, Upturn, 2018.

- A. Bradford, *The Brussels Effect: How the European Union Rules the World*, Faculty Books, 2020, p. 232.
- R. Brauneis – E. P. Goodman, *Algorithmic transparency for the smart city*, Yale Journal of Law & Technology, 2019, p. 104-176.
- F. Bravo, *Data Governance Act and Re-Use of Data in the Public Sector*, European Review of Digital Administration & Law, 2023.
- P. A. E. Brey – J. Soraker, *Philosophy of Computing and Information Technology*, Elsevier, 2009.
- H. Broomfield, *Where Is Open Data in the Open Data Directive?*, Information Polity, 2023, p. 175 – 188.
- T. S. Cabral, *A Short Guide to the Legislative Procedure in the European Union,* EU Law Journal, 2020, p. 161–80.
- T. S. Cabral, *AI and the Right to Explanation: Three Legal Bases under the GDPR*, in (eds.) D. Hallinan – R. Leenes – P. De Hert, Data Protection and Privacy: Data Protection and Artificial Intelligence, Oxford, 2021, p. 29–56.
- E. Carloni, *I principi della legalità algoritmica. Le decisioni automatizzate di fronte al giudice amministrativo*, Diritto amministrativo: rivista trimestrale, 2, 2020, p. 273-304.
- E. Celeste, *Digital constitutionalism: a new systematic theorisation*, International Review of Law, Computers & Technology, 2019, p. 76-99
- Centre for Public Impact, *Destination unknown: Exploring the impact of Artificial Intelligence on Government*, 2017.
- D. Chalmers – G. Davies – G. Monti, *European Union Law: Text and Materials*, Cambridge University Press, 2024.
- D. K. Citron, *Technological Due Process*, Washington Law Review, 2008, p. 1249.
- D. K. Citron – F. Pasquale, *The Scored Society: Due Process for Automated Predictions*, Washington Law Review, 2014, p. 1.
- CJEU, *W. Beus GmbH & Co. v. Hauptzollamt München*, C-5-67, (1968).
- CJEU, *Claude Sayag and S.A. Zurich v. Jean-Pierre Leduc, Denise Thonnon and S.A. La Concorde*, C-9/69, (1969).
- CJEU, *Casper Koelman v. Commission of the EC*, T-575/93, (1996).
- CJUE, *Estabelecimentos Isidoro M. Oliveira SA v. Commission of the EC*, T-73/95, (1997).

- CJEU, *Commission of the European Communities v. Chambre syndicale nationale des entreprises de transport de fonds et valeurs (Sytraval) and Brink's France SARL*, C-367/95 P, (1998).
- CJUE, *Kjell Karlsson and Others*, C-292/97, (2000).
- CJUE, *D e Coster*, C -17/00, (2001).
- CJEU, *Pfizer Animal Health v. Council of the European Union*, T-13/99, (2002).
- CJEU, *Cheil Jedang Corp. v. Commission of the European Communities*, T-220/00, (2003).
- CJEU, *Elf Aquitaine SA v. European Commission*, C-521/09, (2011).
- CJEU, *Otis*, C-199/11, (2012).
- CJEU, *Ziegler v. Commission*, C-439/11, (2013).
- CJUE, *Dano*, C-333/13, (2014).
- CJEU, *Torresi*, C -58/13, (2014).
- CJEU, *European Commission v. Kadi*, C-584/10 P, C-593/10 P and C-595/10, (2014).
- CJUE, *Alimanovic*, C-67/14, (2015).
- CJEU, *Spain v. Commission*, C-521/15, (2017).
- CJEU, *Icap v. Commission*, T-180/15, (2017).
- CJEU, *Sacko,* C-348/16, (2017).
- CJEU, Opinion 1/15 of 26 July 2017.
- CJUE, Case C-73/16, *Puškár*, (2017).
- CJEU, *EG*, C-662/17, (2018).
- CJEU, *Associação Sindical dos Juízes Portugueses*, C-64/16, (2018).
- CJUE, *Achmea*, C-284/16, (2018).
- CJEU, *Commission v. Poland*, C-619/18, (2019).
- CJEU, *R.N.N.S. and K.A. v. Minister van Buitenlandse Zaken*, C-225/19 and C-226/19, (2020).
- CJEU, *La Quadrature du Net u.a.*, C-511/18, (2020).
- CJEU, *Schufa*, C-634/21, (2023).
- J. Cobbe – J. Singh, *Artificial intelligence as a service: Legal responsibilities, liabilities, and policy challenges*, Computer Law & Security Review: The International Journal of Technology Law and Practice, 2021.
- C. Coglianese – E. Lampmann, *Contracting for Algorithmic Accountability*, All Faculty Scholarship, 2021, p. 2311.

- C. Coglianese – D. Ben – M. Lavi, *AI in Adjudication and Administration*, 2021, p. 2118.
- C. Coglianese – D. Lehr, *Regulating by Robot: Administrative Decision Making in the Machine-Learning Era*, Georgetown Law Journal, 17, 2017, p. 147–223.
- J. E. Cohen, *Between Truth and Power: The Legal Constructions of Informational Capitalism*, Oxford University Press, 2019.
- Commission Staff, *Accompanying the document report from the Commission to the European Parliament and the Council On the review of Directive 2016/681 on the use of passenger name record (PNR) data for the prevention, detection, investigation and proesecution of terrorist offences and serious crime*, (July 2020), p. 11.
- Consiglio di Stato, Sec. VI, 8 April 2019, n. 2270.
- Consiglio di Stato, Sec. VI, 13 December 2019, n. 8472.
- Consiglio di Stato, Sec. VI, 13 December 2019, n. 8473.
- Consiglio di Stato, Sec. VI, 13 December 2019, n. 8474.
- Consiglio di Stato, Sec. VI, 04 February 2020, n. 881.
- Consiglio di Stato, Sec. VI, 04 June 2021, n. 1206.
- J. Covelo De Abreu, *The 'Artificial Intelligence Act' Proposal on European e-Justice Domains Through the Lens of User-Focused, User-Friendly and Effective Judicial Protection Principles*, in (eds.) H. S. Antunes – P. M. Freitas – A. L. Oliveira – C. Martins Pereira – E. Vaz de Sequeira – L. Barreto Xavier, *Multidisciplinary Perspectives on Artificial Intelligence and the Law*, Berlin, 2023, p. 397-414.
- P. Craig, *Article 41*, in S. Peers – T. Hervey – J. Kenner – A. Ward, in (eds.) *The EU Charter of Fundamental Rights: A Commentary*, Oxford, 2021.
- P. Craig – G. De Búrca, *EU Law: Text, Cases, and Materials*, Oxford University Press, 2020.
- K. Crawford – J. Schultz, *Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms*, Boston College Law Review, 2014, p. 93.
- K. Crawford – D. Roel – T. Dryer – G. Fried – B. Green – E. Kaziunas – A. Kak – V. Mathur – E. McElroy – A. N. Sánchez – D. Raji – J. L. Rankin – R. Richardson – J. Schultz – S. Myers West – M. Whittaker, *AI Now 2019 Report*, AI Now Institute, 2019.
- J. Danaher, *The Threat of Algocracy: Reality, Resistance and Accommodation*, Philosophy & Technology, 2016, p. 245–68.

- A. Danish – S. Frimpong, *Artificial Intelligence, Machine Learning and Process Automation: Existing Knowledge Frontier and Way Forward for Mining Sector*, Artificial Intelligence Review, 53, 2020, p. 6025-6042.
- G. De Gregorio, *From Constitutional Freedoms to the Power of the Platforms: Protecting Fundamental Rights Online in the Algorithmic Society*, European Journal for Legal Studies, 2018, p. 65.
- G. De Gregorio – S. Demková, *The Constitutional Right to an Effective Remedy in the Digital Age: A Perspective from Europe*, European Yearbook of Constitutional Law 2023: Constitutional Law in the Digital Era, 2024, p. 223-254.
- G. De Gregorio – M. Fasciglione – F. Paolucci – O. Pollicino, *Compliance through Assessing Fundamental Rights: Insights at the Intersections of the European AI Act and the Corporate Sustainability Due Diligence Directive,* MediaLaws, 2024.
- S. De Heer, *Artificial Intelligence and the Right to an Effective Remedy*, in (eds.) A. Quintavalla – J. Temperman, *Artificial Intelligence and Human Rights*, Oxford, 2023, p. 294.
- J. Dirutigliano, *Some considerations on the relationship between the right to a reasoned decision and the right to explanation in the proposal of the artificial intelligence act*, The Digital Constitutionalist the Future of Constitutionalism, 2023.
- District Court of the Hague, 6 March 2020, n. 865.
- C. Ducuing, *The Regulation of Data in the European Union: the Data Governance Act and the Data Act*, (eds.) E. E. Akin – S. Klimbacher – G. Ziccardi, *Smart cities, artificial intelligence and digital transformation law*, Milano University Press, 2024.
- ECtHR, *S. and Marper v. The United Kingdom*, App. 30562/04 and 30566/04, (2008).
- L. Edwards, *Data Protection and Legal Regulation of AI in Europe*, Computer Law & Security Review, 2018.
- H. Eklund, *Article 21*, in (eds.) S. Peers – T. Hervey – J. Kenner – A. Ward, *The EU Charter of Fundamental Rights: A Commentary*, Oxford, 2021, p. 613–638.
- European commission, *Towards a European strategy on business-to-government data sharing for the public interest*, 2020.
- European Commission, *White Paper on Artificial Intelligence - A European approach to excellence and trust*, (February 19th, 2020).

- European Commission, *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*, (February 19th, 2020).
- European Commission, *Coordinated plan on artificial intelligence 2021 review, Annexes to the Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions Fostering a European approach to Artificial Intelligence*, (April 21st, 2021).
- European Commission, *Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions*, A Digital Agenda for Europe (May 19th, 2022).
- European Commission, *Public Buyers Community on the Procurement of AI, EU model contractual AI clauses to pilot in procurements of AI*, (September 29th, 2023).
- European Commission, Joint Research Centre, L. Tangi – A. Rodriguez Müller – M. Combetto et al., *Artificial Intelligence for interoperability in the European public sector – An exploratory study*, Publications Office of the European Union, (October 4th, 2023).
- European Commission, *Public Buyers Community on the Procurement of AI, New version of Procurement Clauses of AI available: supporting responsible use of AI in Public Authorities*, (October 5th, 2023).
- European Commission, *Public Sector Tech Watch latest dataset of selected cases*, 2024.
- European Court of Auditors, *Public Private Partnerships in the EU: Widespread shortcomings and limited benefits*, 2018.
- European Union Agency for Fundamental Rights, *Facial recognition technology: fundamental rights considerations in the context of law enforcement* (November 2019).
- European Union Agency for fundamental rights, *Fundamental Rights Report*, (June 9th, 2020), p. 75.
- S. Feldstein, *The Rise of Digital Repression*, Oxford University Press, 2021.
- A. Ferrario – M. Loi – E. Viganò, *In AI we trust incrementally: A multi-layer model of trust to analyze human-artificial intelligence interactions*, Philosophy & Technology, 2020, p. 523

- M. Fink – M. Finck, *Reasoned A(I)administration: explanation requirements in EU law and the automation of public administration*, European Law Review, 47, 3, 2022, p. 376 – 392.
- L. Floridi, *On the Brussels-Washington Consensus About the Legal Defnition of Artifcial Intelligence*, Philosophy & Technology, p. 1–9.
- L. L. Fuller, *The Morality of Law*, Indiana Law Journal, 1964.
- C. Fusco, *The Use of Artificial Intelligence in the Decision-Making Processes of the Public Administration: Regulations and Executive Practice - The Case of the Italian Public Administration*, in (eds.) D. Marino – M. A. Monaca, *Innovations and Economic and Social Changes due to Artificial Intelligence: The State of the Art, Studies in Systems, Decision and Control*, Berlin, 2023.
- T. S. Gesk – M. Leyer, *Artificial Intelligence in public services: When and why citizens accept its usage*, Government Information Quarterly, 2022, p. 1-12.
- H. Gieskea – I. Van Meerkerk – A. Van Buuren, *The Impact of Innovation and Optimization on Public Sector Performance: Testing the Contribution of Connective, Ambidextrous, and Learning Capabilities*, Public Performance and Management Review, 2019, p. 432-460.
- S. Gilani – A. Al-Matrooshi – M. Khan, *Right of Privacy and the Growing Scope of Artificial Intelligence. Current Trends*, Law and Society, 2023, p. 1-11.
- T. Gillespie, *The Relevance of Algorithms*, in (eds.) T. Gillespie – P. J. Boczkowski – K. A. *Foot, Media Technologies: Essays on Communication, Materiality, and Society*, MIT Press, 2014, p. 167.
- D. Gray – D. Citron, *The Right to Quantitative Privacy*, Minnesota Law Review, 2013, p. 62.
- F. S. Grodzinsky – K. W. Miller – M. J. Wolf, *Developing artificial agents worthy of trust: 'Would you buy a used car from this artificial agent?',* Ethics and Information Technology, 2011, p. 17-21.
- S. Gutwirth – P. De Hert, *Regulating Profiling in a Democratic Constitutional States*, in (eds.) M. Hildebrandt – S. Gutwirth, *Profiling the European Citizen*, 2006, p. 271.
- P. Hacker, *The European AI liability directives – Critique of a half-hearted approach and lessons for the future*, Computer Law & Security Review, 51, 2023.
- H. L. A. Hart, *The Concept of Law*, Oxford University Press, 1961.
- M. Hickok, *Public procurement of artificial intelligence systems: new risks and future proofing*, AI & Society, 2022.

- M. Hildebrandt, *Smart Technologies and the End(s) of Law: Novel Entanglements of Law and Technology*, Edward Elgar Publishing, 2015.
- H. C. H. Hofmann*, Automated Decision-Making (ADM)*, EU Public Law University of Luxembourg Law Research, 2023.
- T. P. Hughes, *Technological Momentum*, in (eds.) M. R. Smith – L. Marx, *Does Technology Drive History? The Dilemma of Technological Determinism*, MIT Press, 1994, p. 112.
- Independent High-Level Expert Group on Artificial Intelligence, *Policy and Investment Recommendations for Trustworthy AI*, (April 8th, 2019).
- C. Intahchomphoo – O. E. Gundersen, *Artificial Intelligence and Race: a Systematic Review, Legal Information Management*, 2020, p. 74-84
- Ipsos, *European enterprise survey on the use of technologies based on AI*, 2020, p. 58.
- J. Laux – S. Wachter – B. Mittelstadt, *Trustworthy Artificial Intelligence and the European Union AI Act: On the Conflation of Trustworthiness and the Acceptability of Risk*, 2022.
- D. Leslie – C. Burr – M. Aitken – J. Cowls – M. Katell – M. Briggs, *Artificial intelligence, human rights, democracy, and the rule of law: a primer*, The Council of Europe and the Alan Touring Institute, 2021.
- J. Locke, *Two Treatises of Government*, 1689.
- A. Jackiewicz, *Prawo do dobrej administracji jako standard europejski (The Right to Good Administration As European Standard)*, Toruń, 2008, p. 58.
- K. Jones, *AI governance and human rights, Resetting the relationship*, Chatman House, 2023.
- S. Junginger, *Transforming Public Services by Design: Re-Orienting Policies, Organizations and Services around People*, Routledge, 2016.
- L. Kaplow, *Rule vs standards: An economical analysis*, Duke Law Journal, 1992, p. 557–629.
- H. Kelsen, *General Theory of Law and State*, Routledge, 1945.
- A. Kesa – T. Kerikmäe, *Artificial Intelligence and the GDPR: Inevitable Nemeses?*, TalTech Journal of European Studies, 10(3), 2020, p. 68-90.
- H. Kissinger – E. Schmidt – D. P. Huttenlocher – S. Schouten, *The age of AI: and our human future*, Little, Brown and Company, 2021.

- L. Koen – M. Kgomotso, *Artificial Intelligence and Racial Discrimination*, in (eds.) J. Temperman – A. Quintavalla, *Artificial Intelligence and Human Rights*, 2023.
- P. D. König, *Fortress Europe 4.0? An Analysis of EU Data Governance Through the Lens of the Resource Regime Concept*, European Policy Analysis, 2022, p. 484–504.
- J. F. M. Koppenjan – B. Enserink, *Public-Private Partnerships in Urban Infrastructures: Reconciling Private Sector Participation and Sustainability*, Public Administration Review, 2009, p. 284.
- G. Krawiec, *Europejskie prawo administracyjne (The European Administrative Law)*, Warszawa, 2009, p. 108.
- M. Lais, *Das Recht auf eine gute Verwaltung unter besonderer Berücksichtigung der Rechtsprechung des Europäischen Gerichtshofs*, Zeitschrift für Europarechtliche Studien, 2002.
- C. Lecher, *What Happens When an Algorithm Cuts Your Health Care*, The Verge, 2018.
- M. Leistner – L. Antoine, *IPR and the Use of Open Data and Data Sharing Initiatives by Public and Private Actors*, 2022.
- J. Lindqvist, *New Challenges to Personal Data Processing Agreements: Is the GDPR Fit to Deal with Contract, Accountability and Liability in a World of the Internet of Things*, International Journal of Law and Information Technology, 2018, p. 45.
- C. Lustig – B. Nardi, *Algorithmic authority: The case of Bitcoin*, 48th Hawaii International Conference on the System Sciences, 2015, p. 743-752.
- J. Maliszewska-Nienartowicz, *Porządek prawny Unii Europejskiej (The Legal Order of the European Union)*, Toruń, 2005, 238.
- N. Marsch, A*rtificial Intelligence and the Fundamental Right to Data Protection: Opening the Door for Technological Innovation and Innovative Protection,* in (eds.), T. Wischmeyer – T. Rademacher, *Regulating Artificial Intelligence*, Berlin, 2020.
- A. C. Martínez*, How can we open the black box of public administration? Transparency and accountability in the use of algorithms*, Revista catalana de dret public, 2019, p. 13-28.
- I. Mendoza – L. A. Bygrave, *The Right Not to Be Subject to Automated Decisions Based on Profiling*, in (eds.) T. Synodinou – P. Jougleux – C. Markou – T. Prastitou, *EU Internet Law: Regulation and Enforcement*, 2017.

- Y. Meneceur, *Artificial Intelligence, Public Administration, and the Rule of Law*, in (eds.) M. Suksi, *The Rule of Law and Automated Decision-Making*, Springer, 2023.
- N. Menéndez González, *The Rights to Privacy and Data Protection and Facial Recognition Technology in the Global North*, in (eds.) A. Quintavalla – J. Temperman, *Artificial Intelligence and Human Rights*, Oxford University Press, 2023, p. 136.
- T. W. Merrill, *The Property Strategy*, University of Pennsylvania Law Review, 160, 2012, p. 35.
- G. Miller – B. Lennett, *White House AI Memo Promises to Balance National Security Interests with Privacy and Human Rights*, Tech Policy.Press, (October 24th, 2024).
- O. Mir Puigpelat, *Algorithms, Automation and Administrative Procedure at EU Level*, University of Luxembourg Law Research Paper, 2023.
- B. Mons – C. Neylon – J. Velterop – M. Dumontier – L. O. B. da Silva Santos – M. D. Wilkinson, *Cloudy, increasingly FAIR; revisiting the FAIR data guiding principles for the European Open Science Cloud*, Information Services & Use, 2017, p. 49-56.
- D. K. Mulligan – K. A. Bamberger, *Procurement As Policy: Administrative Process for Machine Learning*, Berkeley Technology Law Journal, 34, 2019.
- C. O'Neil, *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*, Crown Publishers, 2016.
- P. F. Nemitz, *Constitutional Democracy and Technology in the age of Artificial Intelligence*, Royal Society Philosophical Transactions, 2018.
- J. Niklas, *Human Rights-Based Approach to AI and Algorithms Concerning Welfare Technologies*, in (eds.) W. Barfield, *The Cambridge Handbook of the Law of Algorithms*, 2021, p. 517.
- N. T. Nikolinakos, *The Proposed Artificial Intelligence Act and Subsequent 'Compromise' Proposals: Commission, Council, Parliament*, in (eds.) N. T. Nikolinakos, *EU Policy and Legal Framework for Artificial Intelligence, Robotics and Related Technologies - The AI Act. Law, Governance and Technology Series*, Berlin, 2023, p. 327-741.
- C. Novelli – M. Taddeo – L. Floridi, *Accountability in artificial intelligence: what it is and how it works*, AI & SOCIETY, 2023, p. 1-12.

- T. O'Reilly, *Open Data and Algorithmic Regulation*, in B. Goldstein - L. Dyson, 2013, p. 289-300.
- A. Owusu, *Data sharing in the personal data economy. Does sharing mean caring?*, European Journal of Privacy Law & Technologies, 2023.
- Panel for the Future of Science and Technology EPRS, *The impact of the General Data Protection Regulation (GDPR) on Artificial intelligence*, European Parliamentary Research Service Scientific Foresight Unit (STOA), (June 2020).
- F. Paolucci, *The costs of training AI and the impact on fundamental rights*, The Digital Constitutionalist the Future of Constitutionalism, 2023.
- F. Pasquale, *The black box society: the secret algorithms that control money and information*, Harvard University Press, 2015.
- F. Pasquale, *A Rule of Persons, Not Machines: The Limits of Legal Automation*, George Washington Law Review, 2019, p. 1-54.
- C. Painter, *The UK Coalition government: Constructing public service reform narratives*, Public Policy & Administration, 28, 2013, p. 3–20.
- S. Peers – S. Prechal, Article 52, in (eds.) S. Peers – T. Hervey – J. Kenner – A. Ward, *The EU Charter of Fundamental Rights: A Commentary*, Oxford, 2021, p. 1611–1674.
- W. Pieters, *Explanation and trust: what to tell the user in security and AI?*, Ethics and Information Technology, 2011, p. 53.
- G. Pino, *Il costituzionalismo dei diritti struttura e limiti del costituzionalismo contemporaneo*, il Mulino, 2017.
- A. Pirozzoli, *The Human-centric Perspective in the Regulation of Artificial Intelligence*, European Papers, 2024, p. 105-116.
- K. Pistor, *Statehood in the Digital Age*, Constellations, 2020, p. 3.
- R. Polcák, *Article 12. Transparent Information, Communication and Modalities for the Exercise of the Rights of the Data Subject*, in (eds.) C. Kuner – L. A. Bygrave – C. Docksey, *The EU General Data Protection Regulation (GDPR) A Commentary*, 2020, p 401–402.
- O. Pollicino – G. De Gregorio, *Constitutional law in the algorithmic society*, Cambridge University Press, 2021, p. 3-24.
- J. Ponce, *Good Administration and Administrative Procedures, Indiana Journal of Global Legal Studies*, 2005, p. 567.

- H. Pouget – M. O'Shaughnessy, *Reconciling the U.S. Approach to AI, Carnegie Endowment for International Pease*, (May 3rd, 2023).
- A. Rachovitsa – N. Johann, *The Human Rights Implications of the Use of AI in the Digital Welfare State: Lessons Learned from the Dutch SyRI Case*, Human Rights Law Review, 22 (2), 2022.
- E. Rader – R. Gray, *Understanding user beliefs about algorithmic curation in the Facebook news feed*, Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 2015, p. 173-182.
- C. Ramió Matas, *Inteligencia artificial y administración pública. Robots y humanos compartiendo el servicio público*, Catarata, 2019.
- S. Ranchordas, *The Invisible Citizen in the Digital State: Administrative Law Meets Digital Constitutionalism*, in (eds.) J. De Poorter – C. Oirsouw – G. van der Schyff, *European Yearbook of Constitutional Law*, Tilburg Law School Research Paper, (December 24th, 2023).
- S. Ranchordas, *Experimental regulations for AI: sandboxes for morals and mores*, Morals & Machines 1.1, 2021, p. 86-100.
- J. Raz, *The Authority of Law*, Oxford University Press,1979.
- J. R. Reidenberg, *States and Internet enforcement*, University of Ottawa Law & Techonology Journal, 2004, p. 213.
- D. Reisman – J. Schultz – K. Crawford – M. Whittaker, *Algorithmic Impact Assessments: A Practical Framework for Public Agency Accountability*, AI Now Institute, 2018.
- R. Rodrigues, *Legal and human rights issues of AI: Gaps, challenges and vulnerabilities*, Journal of Responsible Technology, 2020.
- H. Rosenbaum, *Algorithmic neutrality, algorithmic assemblages, and the lifeworld*, 2020.
- J.J. Rousseau, *The Social Contract*, Penguin, 1762.
- J. Ruohonen – S. Mickelsson, *Reflections on the Data Governance Act*, Digital Society, 2023.
- A. Sanchez-Graells, *More Model Contractual AI Clauses – Some Comments on the SCL AI Clauses*, How to crack a nut, 2023.
- A. Sanchez-Graells, *Public Procurement of Artificial Intelligence: Recent Developments and Remaining Challenges in EU Law*, Legal Tech Journal, 2024.

- A. Sanchez-Graells, *Resh (AI) ping Good Administration: Addressing the mass effects of public sector digitalization*, Laws, 2024, p. 9.
- G. Sartor, *Artificial Intelligence and Human Rights: Between Law and Ethics*, Maastricht Journal of European and Comparative Law, 27, 2020, p. 705–719.
- G. Sartor, *Comparative Constitutional Engineering*, 1994.
- C. Scarpellino, *EU and US regulatory approach to AI: a comparative perspective*, LUISS Policy Observatory, 2024.
- K. Shirley – S. Ranchordas – S. Van De Wetering, *AI failure, AI success, and AI power dynamics in the public sector*, 2024.
- C. Schmitt, *Constitutional Theory*, Duke University Press, 1928.
- G. Schneider, *The Algorithmic Governance of Administrative Decision-Making: Towards an Integrated European Framework for Public Accountability*, Eurojus, 2019, p. 134-148.
- B. Schölkopf, *Causality for Machine Learning*, in (eds.) H. Geffner – R. Dechter – J.Y. Halpern, Probabilistic and Causal Inference, New York, 2022, p. 765–804.
- N. Seaver, *Knowing algorithms*, Media in Transition, 2013, p. 1-12.
- Secretary General, *Report of the Special Rapporteur on the promotion and protection of the rights to freedom of opinion and expression*, (August 18th, 2017).
- A. Simoncini, *Amministrazione digitale algoritmica. Il quadro costituzionale*, Il diritto dell'amministrazione pubblica digitale, 2020.
- A. Simoncini – E. Longo, *Fundamental rights and the rule of law in the algorithmic society. Constitutional challenges in the algorithmic society*, Cambridge University Press, 2021, p. 27-33.
- N. Seaver, *Algorithms as culture: Some tactics for the ethnography of algorithmic systems*, Big Data & Society, 2017, p. 1-12.
- M. Szydło, *Prawo do dobrej administracji jako prawo podstawowe w unijnym porządku prawnym (The Right to Good Administration as Fundamental Right in the Union's Legal Order)*, Studia Europejskie, 2004, p. 95.
- M. Tajabadi – L. Grabenhenrich – A. Ribeiro – M. Leyer – D. Heider, *Sharing Data With Shared Benefits: Artificial Intelligence Perspective*, Journal of Medical Internet Research, 2023.
- L. Tangi – C. van Noordt – M. Combetto – D. Gattwinkel – F. Pignatelli, *AI Watch. European Landscape on the Use of Artificial Intelligence by the Public Sector,* Publications Office of the European Union, 2022.

- Tar Lazio, sec. III-*bis*, 10 October 2018, n. 9224.
- Tar Lazio, sec. III-*bis*, 10 September 2019, n. 10964.
- D. Tatevik, *The U.S. Approach to AI Regulation: Federal Laws, Policies, and Strategies Explained*, (September 09th, 2024).
- C. Thönnes – N. Vavoula, *Automated predictive threat detection after Ligue des Droits Humains: Implications for ETIAS and CSAM (Part I)*, VerfBlog, 2023.
- D. Tiberiu – Y. Lupu, *Digital authoritarianism and the future of human rights*, International Organization, 2021, p. 991-1017.
- L. Torchia, *Lo stato digitale*, Il Mulino, 2023.
- R. Van Den Hoven Van Genderen, *Privacy and Data Protection in the Age of Pervasive Technologies in AI and Robotics*, European Data Protection Law Review, 3, 2017, p. 338 – 352.
- L. Van Wichelen – J. C. Devogelaere, *Artificial Intelligence: From public discrimination to public administration*, The EUTOPIA Student Think Tank (EUSTT), 2023.
- L. Vandercruysse – A. Christofi – C. Buts – M. Dooms – P. Valcke, *Data Protection in Smart Cities*, European Procurement & Public Private Partnership Law Review, 2022, p. 81 – 93.
- A. Venn Dicey, *Introduction to the Study of the Law of the Constitution*, Roger E. Michener, 1885.
- J. Vested-Hansen*, Article 7 (Private Life, Home and Communications)*, in (eds.) S. Peers – T. Hervey – J. Kenner – A. Ward, *The EU Charter of Fundamental Rights: A Commentary*, Oxford, 2021, p. 151–194
- A. Voorwinden, *The privatised city: technology and public-private partnerships in the smart city*, Law, Innovation and Technology, 13, 2, 2021, p. 439-463.
- S. Wachter – B. Mittelstadt – L. Floridi, *Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation*, International Data Privacy Law, 2017.
- A. M. Walorska, *The Algorithmic Society*, in (eds.) D. Feldner, *Redesigning Organizations Concepts for the Connected Society*, Springer, 2020.
- M. Weber, *Economy and Society*, Bedminster Press, 1922.
- F. Webster, *Theories of the Information Society*, Routledge, 2014.
- N. Wiener, *The Human Use of Human Beings: Cybernetics and Society*, Da Capo Press, 1988.

- M. Willson, *Algorithms (and the) everyday*, Information, Communication & Society, 2017, p. 137-150.
- M. Wörsdörfer, *Biden's Executive Order on AI and the E.U.'s AI Act: A Comparative Computer-Ethical Analysis*, Philosophy & technology, 2024, p. 74.
- WP29, *Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679*, (February 6th, 2018).
- K. Yeung, *Algorithmic Regulation: a Critical Interrogation, Regulation & Governance*, 2017, p. 20.
- M. Zalnieriute – L. B. Moses – G. Williams, *The Rule of Law and Automation of Government Decision-Making*, Modern Law Review, 2019, p. 245.
- M. Zalnieriute – L. B. Crawford – J. Boughey – L. B. Moses – S. Logan, *The Cambridge Handbook on the Law of Algorithms*, Cambridge University Press, 2019, p. 30.
- G. Zanfir-Fortuna, *Four Data Protection Threads in Today's Biden-Harris EO on AI, through a Global lens,* (October 30th, 2023).
- H. Zech, *Liability for AI: public policy considerations, ERA Forum*, Journal of the Academy of European Law, 2021, p. 147 – 150.
- S. Zuboff, *The Age of Surveillance Capitalism. The Fight for a Human Future and the New Frontier of Powe*r, Public Affairs, 2019.