FINAL DEGREE PROJECT

Artificial Intelligence and Public Administration:
Ethical Implications of the Use of Automatic Decision-making Systems

AUTHOR: Francesca Toniolo

TUTOR: Daniel Sansó-Rubert Pascual

GRADO EN RELACIONES INTERNACIONALES/
BA IN INTERNATIONAL RELATIONS

Academic year 2022/2023

FACULTAD DE CIENCIAS SOCIALES Y DE LA COMUNICACIÓN

UNIVERSIDAD EUROPEA DE MADRID

ABSTRACT

As new technologies developed to the point in which their presence reaches all sectors of public and private life, so too have concerns about their associated ethical impacts. Although alarms about the application of Artificial Intelligence (AI) to public administration started to arise in the last decades, the topic has yet to obtain particular attention. This paper aims at the analysis of existing legislation regarding the use of AI in the European Union and present different cases of application of automated decision-making (ADM) systems in national administration to assess the role of ethics in these processes. Furthermore, the analysis argues the importance of the system's adhesion to ethical values for the promotion and preservation of citizens' well-being and dignity.

KEYWORDS:

RESUMEN

A medida que las nuevas tecnologías se han desarrollado hasta el punto de que su presencia alcanza todos los sectores de la vida pública y privada, también lo ha hecho la preocupación por sus repercusiones éticas asociadas. Aunque las alarmas sobre la aplicación de la Inteligencia Artificial (IA) a la administración pública comenzaron a surgir en las últimas décadas, el tema aún no ha obtenido especial atención. El objetivo de este artículo es analizar la legislación existente sobre el uso de la IA en la Unión Europea y presentar diferentes casos de aplicación de sistemas de toma de decisiones automatizada (ADM) en la administración nacional para evaluar el papel de la ética en estos procesos. Además, el análisis argumenta la importancia de la adhesión del sistema a los valores éticos para la promoción y preservación del bienestar y la dignidad de los ciudadanos.

PALABRAS CLAVES:

Inteligencia Artificial, Sistemas Automatizados de Toma de Decisiones, Ética, Buen Gobierno, Administración Pública, Europa, Legislación.

INDEX OF ACRONYMS AND ABBREVIATIONS

ACRONYM	ENGLISH	SPANISH
ADM	Automated decision-making	Toma de decisiones automatizadas
AI	Artificial Intelligence	Inteligencia Artificial
BKA	German Federal Criminal Police Office (Bundeskriminalamt)	Oficina Federal de Policía Criminal de Alemania
CNMC	National Commission on Markets and Competition	Comisión Nacional de los Mercados y la Competencia
ECHR	European Convention on Human Rights	Convención Europea de los Derechos Humanos
EEA	European Economic Area	Espacio Económico Europeo
EGDI	United Nations E-Government Development Index	Índice de Desarrollo del Gobierno Electrónico
EU	European Union	Unión Europea
FNV	Dutch Federation of Trade Union (Federatie Nederlandse Vakbeweging)	Federación Holandesa de Sindicato
GDPR	General Data Protection Regulation	Reglamento General de Protección de Datos (RGPD)
IMF	International Monetary Fund	Fondo Monetario Internacional
IO	International Organization	Organización Internacional
MIUR	Italian Minister of Education (Ministero dell'Istruzione e del Merito)	Ministerio Italiano de Educación
NGO	Non-Governmental Organization	Organización non-gubernamental
NJCM	Netherlands Committee of Jurists for Human Rights (Nederlands Juristen Comité voor de Mensenrechten)	Comité de Abogados Holandeses por los Derechos Humanos
NPG	New Public Governance	Nueva Gobernanza Pública
NPM	New Public Management	Nuevo Gerencialismo Público
OHCHR	United Nations Office of the High Commissioner of Human Rights	Oficina del Alto Comisionado de las Naciones Unidas para los Derechos Humanos

PVPC	Voluntary Price for the Small Consumer	Precio Voluntario para el Pequeño Consumidor
RADAR-iTE	Rule-based analysis of potentially destructive perpetrators to assess acute risk – Islamist terrorism	Análisis basado en reglas de autores potencialmente destructivos para evaluar el riesgo agudo - Terrorismo islamista
RISKANT	Risk analysis of those inclined to act on Islamist motivations	Análisis del riesgo de los proclives a actuar por motivaciones islamistas
SUWI Act	Work and Income Structure Implementing Organization Act	Ley orgánica de aplicación de la estructura de trabajo e ingresos
SZW	Ministry of Social Affairs and Employment (Ministerie van Sociale Zaken en Werkgelegenheid)	Ministerio de Asuntos Sociales y Empleo
SyRI	System Risk Indicator (Systeem Risico Indicatie)	Indicador de Riesgo del Sistema
TAR	Regional Administrative Court (Tribunale Amministrativo Regionale)	Tribunal Administrativo Regional
UN	United Nations Organization	Organización de las Naciones Unidas (ONU)
UN DESA	United Nations Department of Economic and Social Affairs	Departamento de Asuntos Económicos y Sociales de las Naciones Unidas
UNDP	United Nations Development Programme	Programa De Las Naciones Unidas Para El Desarrollo (PDNU)

TABLE OF CONTENTS

1. INTRODUCTION

The wheel, the plow, numbers, writing, the light bulb, the telescope, the microscope, printing, and computers are great symbols of technological advances that changed history. They can be considered new forms to see the reality that modified and radically affected people's lives. Today we live in a world dominated by technology; we are experiencing new forms of development and change, achieving the point at which machines can replicate human intellectual processes and perform various advanced functions. Artificial Intelligence (AI) systems have been in the past decades developing and expanding. Used at first only to play chess, they reached competencies that allowed them to be introduced in most sectors of society, including the public administration sector.

This development has been highly taken advantage of in Europe, especially inside the European Union, which following its aspiration of development, highly promoted digitalization in all sectors of society. In this regard, as of 2022, Europe is the continent that, according to the United Nations E-Government Development Index (EGDI)[1], has the highest average EGDI value, representing progress and technology integration in governmental processes. With an EGDI corresponding to 0.8602, Europe highly detaches from Asia (0.6493), Americas (0.6438), Oceania (0.5081), and Africa (0.4054) (UN DESA, 2022). The progress of technology, demonstrated by the ease of information, content, and service delivery, unavoidably brought technological inputs in core governance processes, such as decision-making.

In an era marked by rapid technological advancements, automated decision-making (ADM) systems have been introduced in all sectors of society, including public administration. These systems, driven by AI and machine learning algorithms, hold the potential to improve processes enhancing efficiency and service delivery while at the same time raising concerns about their ethical implications, especially when citizens are affected. Given the rapid development of such systems, the ease with which they are applied in all sectors of society, and the blind spots of their application, it is of fundamental importance to drive more attention to the issue. In accordance, this

---

[1] As indicated in the 12th edition of the UN E-Government Survey 2022, published by the UN Department of Economic and Social Affairs (UN DESA).

investigation aims explicitly at the analysis of the ethical impact of algorithmically driven automated decision-making (ADM) systems on citizens in public administration.

In the enactment of modern democratic governance, ethics is a fundamental factor to take into consideration when assessing the correctness of authority toward citizens. According to the European Commission, practitioners of public administration should be driven by ethics, translated into the application of ethical rules and values that cannot be altered or bypassed to ensure that the public is treated fairly and equitably for the perpetration of good governance (European Commission. Directorate General for Employment, Social Affairs and Inclusion, 2017).

However, with the introduction of automated systems and AI in governmental processes, the consistency and perpetuity of these essential factors are endangered as the actors central to the decision-making process mutate. Because of this possible change, it is of fundamental importance to assess whether ethics and accountability can continue to be perpetrated when AI technologies are inserted into public governance processes to create automated responses that can profoundly influence the lives of citizens.

While technologies such as AI and ADMs have the potential to be highly convenient and beneficial for human beings, their inappropriate use or inopportune programming can cause intense damage. It is therefore essential that their introduction in such fields, which have a tangible and inevitable impact on the lives of people, as in the case of public governance and, more generally, the public sectors, is correctly directed by specific legislation, controlled, and audited to function in accordance with the ethical values of good governance.

This paper will analyze the application of software-based systems, considering those used in automated responses and decision-making processes in the realm of public administration. The area of study under examination will be the European Union, processed with a general analysis of existing legal regulations on AI in the region and a specific analysis of the application of an ADM system in the public governance of four different EU countries. The focus of this research will be limited to the ethical and governmental perspective of the use of AI, omitting the deepening of the analysis

under the technical and engineering lens of the design, development, and use of AI systems.

1.1 Objective

To investigate the ethical impact AI has on citizens when applied to governance, the object of study of this analysis is the specific application of ADM systems in decision-making processes in public administration. This object has been selected to restrain the research and focus on processes that were once developed by human beings and are now completed by algorithmic-based systems that operate following an input-output scheme based on given data. Since ADMs are the most diffused type of AI implemented in public governance and specifically in public administration's decision-making, a sector that directly affects the lives of citizens, their analysis allows the correct assessment of the impacts of the digitalization of public administration.

The study has been conducted to identify whether using ADM systems in public administration has a positive, negative, or neutral impact from an ethical standpoint for citizens. Considering that ADMs, as AI, are key representatives of the development and progress of society as new technologies and that their application is inevitable in all sectors of public and private life, it is important to evaluate their impact. Hence, it becomes crucial to estimate if the type of technologies that it is dealt with, at the current point of evolution, can be safely and ethically implemented. Furthermore, it must ensure that their deployment aligns with the principles of good governance or jeopardizes the life of civilians.

As the topic of analysis is a highly contemporary issue subject to continuous investigation and discussion, that still embodies various questions for the public and academia; this study aims to draw attention to the highly diffused and unexposed use of ADMs, specifically in the field of public administration. Furthermore, it has the objective of examining legal and regulatory frameworks related to the use of AI present in the European Union. At the same time, it highlights the importance of ethical consideration prior to the implementation of algorithmic systems that substitute human beings in decision-making processes. Finally, the research intends to evaluate specific existing ADM systems used in public administration, assessing their compliance with

ethical principles and underlining how distancing from those principles impacts the lives of citizens or specific vulnerable groups.

1.2 Methodology

This analysis follows a qualitative model of investigation to explore the use of automated decision-making systems in public governance within the European Union. Given the recentness and the continuous evolution of the topic of study and the purpose of the research, which aims at the analysis of specific ethical impacts, the data taken into consideration are of qualitative nature rather than a quantitative one. Data has been collected through the conduction of an extensive literature review and the observation of four specific case studies. The study draws on primary and secondary data sources, encompassing the revision of specialized bibliographies such as academic journals, official documents, reports, and legislative texts.

Using a deductive style, the analysis tackles the topic from the general to the specific, from an international to a national perspective, and provides an in-depth study on the general application of ADM systems in Europe and the existing EU legislation which interests the international sphere. This is followed by the comparison of national perspective with an exemplification of four relevant cases of diverse applications of ADMs in the public administrations of four different countries of the European Union, namely the Netherlands, Italy, Spain, and Germany.

The work has been structured into five different sections. The first section, (1) *Introduction*, is dedicated to a general introduction to the topic of study, which justifies the selection of the subject of research, the objectives of the investigation, and the methodology with which the investigation has been conducted. The second part, (2) *Theoretical framework,* corresponds to the analysis of the theoretical framework in which the study is developed, which comprises of the application of international relations and sociological theories to the concepts of good governance, ethics, and the conceptualization of the application of automated algorithmic systems to decision-making. The third section, (3) *Conceptual Background*, includes the conceptualization of terms used through the analysis of which the understanding is fundamental for the correct interpretation of the study, those being AI and ADM, Good Governance, and Ethics. The fourth chapter, (4) *Research Development,* corresponds to the actual

development of the study, divided into six subsections it will tackle: (4.1) *Legislative Overview of ADMs Application in European public administration*, an overall analysis of the laws and prescriptions behind the general application of AI in public administration in Europe; followed by the exemplification of four (4.2) Case studies: (I) *The case of SyRI, ADMs for fraud detection in the Netherlands, an extensive* data analysis system used by the Dutch government to store and compare citizens' data between different public institutions; (II) *The case of Buona Scuola, ADMs for the allocation of professional destinations in Italy*, a system used by the Italian government to reorganize the national public school system by assigning new locations to school teaching staff; (III) *The case of BOSCO, ADMs for the distribution of financial aid in Spain*, tackling the case of an ADM system integrated into Spanish public administration for the allocation of public energetic financial aid; and (IV) *The case of RADAR-iTE, ADMs for risk assessment and recognition of possible terroristic threats in Germany*, a system used by German authorities in the field of state security to prudently prevent outcomes of radicalization in the country. In conclusion to this section, (4.3) *Similarities and differences,* include the final comparison between the proposed examples and a judgment of the ethical impacts these cases had on national civilians. Finally, the analysis resolves with a (5) *Conclusion,* which includes an assessment of the results found in the study and some final reflections on the topic.

The study's design and methodological approach are intended to thoroughly examine automated decision-making systems' appliances in public administration at the European Union level. The goal is to obtain an exhaustive comprehension of this phenomenon through a comprehensive analysis. The study draws on primary and secondary data sources and employs a comparative case study method to identify recurring themes and patterns in the data. It is noted that there exist certain limitations within the scope of this study, therefore a comprehensive understanding and appraisal of the outcomes must consider how the objective of the study has been framed and the difficulties encountered in the elaboration of the study because of the continuous and rapid evolution of the topic of study and its relative regulation.

2. THEORETICAL CONCEPTUAL FRAMEWORK

Since Greek times the question of ethics has been discussed as a transcendental factor of governance; according to the Athenian philosopher Thucydides, ethics is the "philosophy of morality". Instead of seeing morality as a universal, natural force, the author sees it as the effect caused by a person's physical and social environment. Following this idea, Thucydides saw ethics as the study of various factors in a person's life that influence their morality.

In his prominent work *History of the Peloponnesian War* (431 B.C), Thucydides dedicated the first chapter of the third book, Chapter X, to the description of the Corcyraean revolution, as a representation of the collapse of civilization in which family structures and moral law are two of the bonds that kept society alive. Thucydides articulates that a stable, functional, and effective overarching institution overseeing human instincts is necessary for creating moral behavior in society (Thucydides, 2003). In Thucydides' idea, the state's effectiveness plays a significant role in controlling and directing human instincts, creating social morale. Following such a realist perspective, individuals driven by self-interest would behave morally only under state coercion.

Consequently, Thucydides reflects on how the ethics of individuals in authoritative positions affect their governance and their relationship with subordinates. Specifically, this reflection can be observed in his assessments of the authoritarian figures of Pericles and Alcibiades (Mantzouranis, 2018). In them, Thucydides respectively highlighted how the prioritization of principles of the common good in Pericles' government affected civil life, creating societal benefits, while Alcibiades' preference towards personal aggrandizement and economic enrichment represented a potential threat to the best interests of society.

With the comparison between these two statesmanlike figures, Thucydides suggested understanding leadership as a relational process, a notion predecessor to the conceptualization of the Authentic Leadership Theory (Mantzouranis, 2018). Representing an essential current of modern leadership that conceive ethics and effectiveness as two fundamental components for the conception of good leadership, the Authentic Leadership Theory emphasizes the importance of a leader's values,

ethics, and actions and the relative transparency and consistency with which these are deployed to guarantee accountability at the same time as discretion (Kellis & Ran, 2013). Authentic leadership aligns with the principles of modern democracy, and its application should be considered for the conduction of governance based on the principles of ethics to achieve good governance.

Thucydides' ideas have been shared by two great exponents of ancient Greek philosophy, Plato (427- 347 BC) and Aristotle (384-322 BC), who shaped the Western intellectual tradition and the history of Western philosophy. In ancient Athens, democracy was based on a social contract in which citizens had a great say in, and responsibility for the decisions made. In his book *The Republic*, Plato exposes how the human being and the state share three fundamental characteristics which make them highly similar. Reason, Spirit, and Appetite were demonstrated as characteristics of each individual and impersonated in the roles of the Philosopher, the Soldier, and the Workers in the state. Each of the state's roles reflects one of the individual's properties, used as tools for the realization of the state, which needs the presence of government, protection, and the flow of the economy. This division and specialization of roles for the correct function of the state avoid moral corruption and guarantee the promotion of good life or goodness (Basak & Chakroborty, 2018).

While Plato depicted the necessary factor for creating a state in which the Philosopher King, as the only personality rooted in values, conducts the state and creates the government; Aristotle lays the basis for the creation of the concept of *good governance* and the ways to achieve it. He suggested the implementation of *Polity* as a form of governance that blends the extremes of oligarchy and direct democracy and tends more toward democracy, with ethical and social foundations and a civilization predominantly composed of members of the middle class. Furthermore, he notably underlined the importance of the rule of law and the implementation of virtue by citizens and civil servants. According to Aristotle, worthy administration primarily aims to familiarize citizens and civil workers with virtue (Basak & Chakroborty, 2018). Ethical dilemmas in public governance, or decision-making procedures, need to be guided by an ethical compass that guarantees the democratic answerability of governance, the notion of legality and the application of the law, official honesty, and alertness to civil society's needs.

From this perspective, ethics analysis is fundamental for correctly applying good governance. The theoretical framework of virtue ethics is one of the most antique approaches to ethics, strictly linked to Aristotle and subsequently embraced by scholars such as Professor Alasdair Macintyre (Lynch, 2004). Conceiving everything in terms of material, efficient, formal, and final causes (in ancient Greek: *telos*)*,* Aristotle believed in the achievement of *the good life* as the *telos* of each individual, which the master of intellectual virtue can achieve as the control of human desires done by one's moral virtue. In the twentieth century, Macintyre experienced the urge to restore the lost Aristotle's theory of ethical virtue, understanding everything from a functionality conception and therefore introducing rationality into the moral debate, denoting the importance of the social contextual nature of human existence. The theory of virtue ethics requires the exercise of virtues and encourages the development of ethical judgments in its members. According to Macintyre's vision, the satisfaction of virtue in public administration for the achievement of its *telos,* the benevolent pursuit of the public interest, provides both an internal and external good (Lynch, 2004).

In the eighteen century, utilitarianism emerged as a novel theory in the field of normative ethics by the hand of Jeremy Bentham. The utilitarian theory was created as a specific strand of consequentialist ethics that affects public administration in decision-making, focusing on optimizing benefits in the outcome and result of perpetrated actions. According to Bentham, an action should be approved to the extent to which its direct consequences increase the happiness of the party whose interests are in question and disapproved when the consequences oppose that happiness. In this sense, utility is considered an action's final end or consequence. Specifically, *"an action is conformable to utility when the tendency it has to augment the happiness of the community is greater than any it has to diminish it"* (Bentham et al., 1996: 12-13). In governance, utilitarianism predilects the interests of society over the individual rights of leaders, going against the theory of social contract. Despite its importance in current studies, it remains an abstract ethical doctrine in the field of normative ethical philosophy because of its conceptual limitations and lack of practicability, given the difficulties one can encounter in distinguishing a specific action in a decision-making process and assessing all the interlinking, practical and abstract consequences of an action.

The intersection between utilitarianism and liberalism, a current promoting equality, the protection of human rights, and individual liberty for the final aim to assure the well-being of citizens, necessarily procured a change in the theory itself. Indeed, as a complement to Benthamite's theory, John Stuart Mill opted for the insertion of liberal principles in classical utilitarianism to form what can be defined as the theory of liberal utilitarianism. This supplement allowed the evolution of the theory toward a more human rights-centered and liberal theory, framed so that the prevention of harm to others, created by one's action or inaction, is the sole purpose for which power can be rightfully exercised over members of a civilized community (Häyry, 1994).

In contrast to the utilitarian theory, the deontology ethical perspective affirms that actions are ethical if they align with a specific set of rules; exclusively satisfying one's duties, ethical practices are endorsed. Deontology, meaning the study of duty, affirms the impossibility of judging all actions by their consequences, underlining the possibility of an action producing a positive outcome but not conforming with moral norms. Deontology plays a vital role in public administration, as it involves fulfilling the needs of the people and acting as a link between law enforcement and ethics. Under this ethical conception, public officers are legally obligated to serve society and its citizens while adhering to moral principles that define their profession. Although this concept originated from the philosopher Immanuel Kant's belief in treating humanity as an end rather than a means, a modern approach to deontology would require a more comprehensive evaluation of risks, moral considerations, and consequences when applied to governance (Larry & Moore, 2021).

The conception of a value-based public administration contrasts with a pure governmental administration dominated by the business-oriented values of economy, efficiency, and effectiveness. Hence, according to Professor Christopher Hood, the distinction between a solely value-based administration and a solely business-based administration needs to be remedied by complementing ethical values and business aims for good governance (Basak & Chakroborty, 2018).

Hood promoted the theory of New Public Management (NPM) as both a social movement and the subject of academic study created in the 1980s to denote the importance of introducing principles and practices of the private sector into the public sector. Following Hood's ideas, this theory aims to deliver more attention to the

satisfaction of citizens' demands and service delivery through organizations other than traditional public bureaucracies (Hood, 1995). In this sense, NPM is based on principles of democratic governance, and it deploys results-driven actions with an entrepreneurial attitude that aims at people empowerment, service decentralization, and accountability.

Notwithstanding NPM remains one of the dominant approaches in public administration; changes such as the rise of globalization, the introduction of new data-based technologies, and the evolution of communication and information technologies produced a new public administration theory, denominated New Public Governance (NPG). The emergence of NPG represented the restructuring and amplification of the scope of the NPM and the development of a more holistic model of public policy. Indeed, NPG indeed maintained the principles of NPM, allowing the implementation of a new form of horizontal governance that endorsed a public-private partnership that included the participation of new actors in the administration of public governance, contributing to the creation of a multifaceted social context that could better respond to civilian needs (Andrikopoulos & Ifanti, 2020).

The era of NPM, which dominated the 1990s, was replaced in the first decade of the 2000s by a new form of public administration ideology, Digital-era Governance. As Professor Patrick John Dunleavy and his colleagues stipulated, this era represents the extensive digitalization of administrative operations, bringing an infrastructural, administrative, and cultural shift in public governance (Dunleavy et al., 2005). Initially related to the introduction of information and communication technologies, it evolved with the rise of new technologies. In it, technologies started to play a relevant role in public administration to deliver better, faster, and more efficient communication, which invoked financial, administrative, and managerial considerations.

The rise of a new era of digitalization, brought about by the introduction of new digital technologies, disclosed a new set of public service issues, including automation, decentralization, and dehumanization, touching upon considerations about ethical and moral challenges. The difference between the application of initial digital technologies and the new digital technologies in public administration, a part of laying in the apparent progress of the digital tool used, stays in the aim with which these tools are used. Indeed, rather than the mere transfer of information, new technologies permit

the acquisition of information from the exploitation of data in the public and private domain (Tan & Crompvoets, 2022).

While in the former Digital-era Governance, the aim was to allow real-time delivery of information, with the rise of the New Digital-era Governance, the objective shifted from faster data to better data for the creation of citizens-centered services. This shift created various problematic implications regarding social and political trust and value creation in public service provision. Regarding trust, trust issues are shaped by the invasive use of technologies and data analytics' predictive and manipulative power on users and behaviors. In the form of value creation, new technologies add value to administrative processes, increasing the efficiency of service delivery. However, their use can corrode the created value, unveiling challenges in public accountability (Tan & Crompvoets, 2022).

This corrosion can be expressed in various forms, for example, by not preventing discriminatory policy practices, increasing political failure costs, or lacking an effective system of control and auditing that protects citizens against technological abuses. Because of the problems created by the advent of new technologies in public administration, their introduction requires, among other transformations, a restructuring of the organizational role of staff and the creation of new positions that specialize in data and digital technologies (Tan & Crompvoets, 2022).

To better understand how new technologies can interact with and affect the process of governance in public administration theoretically, it is possible to refer to French philosopher Michel Foucault's concept of *governmentality*. The theory of governmentality initially developed by Foucault between the 1970s and 1980s, similar in essence to Aristotle's understanding of good governance, provides essential insights into how power is exercised through institutions and practices and can deliver an efficient understanding of its functioning in contemporary societies.

With the term governmentality Foucault englobed the two distinct concepts of *government*, which refers to the governance of conducts or human behaviors, and *rationality*, the attempt to bring any form of rationality into the acts of governance (Dean, 2010). Including both the practice of conduct of oneself and the conduct of others, governmentality deals with the reasoning that stays behind governing, as any

act of decision-making, in which systematic ways of reasoning are preferred over other models. Governmentality emerged in the early modern era in Western European societies, coinciding with the rise of a liberal form of government focused on the population, its health, welfare, prosperity, and happiness, and engaging in the emergence of a political economy.

Different forms of governmentality have been dominating in different periods, creating the three families of governmentality by Miller and Rose, the *Classical liberalism* of the Nineteen Century, the *Governing from the social point of view* of the early Twenty Century, and the *Advanced liberalism* of the late Twenty Century, in which different governmental conducts and circumstances characterize each one. The change in society that increased the use of digital networks and algorithms of the twenty-first century led to the rise of a fourth family of governmentality in which technologies facilitated a change in the relationship between administrative authorities and citizens (Dent, 2020).

New freedoms, new social relations, and new cognitive orientations give rise to *smart cities* where information is at everyone's hand, and bureaucratization is expanded rather than reduced, becoming universal with the proliferation of procedures, new forms of surveillance, and rulings. This universal bureaucratization reserves ideas of governmentality, with the conjunction between conduct and rationality, but as well reflects the neoliberal Weberian vision of bureaucracy rooted in efficiency, objectivity, and rationality. In this prospect, Weber envisions efficiency as a measurement of the suitability of means to achieve a specific goal, objectivity as the dehumanization of processes and exclusion of human sentiments, and rationality as the sociocultural manifestation of capitalistic reason. Weber's outline of bureaucracy is highly influenced by the historical period in which the theory was shaped, reflecting in all its parts principles of industrial capitalism (Muellerleile & Robertson, 2018).

Unlike the capitalistic industrialization of the twentieth century, technology has now impregnated both capitalistic production and social realities, becoming more present in every sector of society. In the digitalization of bureaucratic processes, Weber's principles of efficiency, objectivity, and rationality endure, with a difference in the creation of a new sort of rationality that distances itself from formal rationality thanks to the augmentation of information reach (Muellerleile & Robertson, 2018). From this

perspective, it can be acknowledged the inevitability of technological inclusion in governmental processes. Underlining the necessity of dedicated accouterments for protecting ethical values and the perpetration of good governance.

From a more pessimistic point of view, José Maria Lasalle sees the rise of technology as resulting in the collapse of liberal democracy. In his book *Ciberleviatán: el colapso de la democracia liberal frente a la revolución digital,* Lasalle identifies the digital revolution as a progressive and evolutive catastrophe in which the technical, political, and economic align in favor of the elite, generally hindering the rest of society. However, the author sees the imposition of a *Technological Big Deal*, a pact between technology and the human being in which humanism prevails over the power of technology, as the only possible instrument for a liberal uprising (Lassalle, 2019). This theory is shared by the Italian philosopher and scholar Luciano Floridi, who sustains that technology undoubtedly has high ethical implications for human enhancement. It is only with the responsible deployment of AI, in which ethical considerations are made from the start of its design, that the promotion of human well-being and respect for human values can be assured (Floridi, 2013).

To the formation of Lasalle's deal, it is necessary to put the digital transformation of public administration under ethical debate. In the framework of data-driven urbanism, which can be applied to our field of research, Ph.D. Rob Kitchin developed two categories to classify ethical concerns regarding the involvement of data in governance. The first one, denominated *datification and privacy*, comprises all those privacy concerns, tackling the sphere of identity, body, territory, location, communication, and transactions, which brings the author to stress the danger of the recent intensification of surveillance and interference in human life, perpetrated without given consent. While the second category, named *data use, sharing, and repurposing*, refers to the current lack of data minimization, which leads to the creation and storage of gratuitous data that can then be repurposed. This repurposing leads to concerns about *data determinism*, a practice used in predictor governance to anticipate citizens' behaviors and actions, for example, to predict possible threats (Kitchin, 2016).

From a similar perspective, in the application of technology to governance and specifically decision-making, Professor David Bianchini and Ismael Ávila doubt the

widely granted openness and social acceptance of patterns used, remarking their complexity and often blurriness, which lead to the necessity to root processes in ethical values and use them as parameters for their domination. Furthermore, they reflect in their article *Smart Cities and Their Smart Decisions: Ethical Considerations* on the lack of neutrality of technology, which in the absence of appropriate control, can differently treat opposite social classes falling into the discrimination of vulnerable groups (Bianchini & Avila, 2014). The presence of biases in AI systems, as underlined by other authors such as Professor Michael Kearns and Professor Aaron Roth, reflects and often amplifies the biases present in the data with which systems are trained. Being, therefore, not inherently present in the algorithm-based systems makes AI systems amendable, in the case in which the bias can be identified, when aligned with societal values and fed by a diversified representative data (Kearns & Roth, 2020).

3. CONCEPTUAL BACKGROUND

3.1 Artificial Intelligence and Automated Decision-making Systems

The term Artificial Intelligence (AI) was coined in 1955 by the emeritus Stanford Professor John McCarthy and was officially presented the following year at the Dartmouth Summer Research Project on Artificial Intelligence, regarded as the official event marking the birth of the research field (Moor, 2006). McCarthy defined the term as *"The science and engineering of making intelligent machines, especially intelligent computer programs"* (McCarthy, 2007: 2)[2]. From his vision and belief in AI's potentiality, intelligent machines' evolution brought systems and, therefore, core algorithms to evolve toward the actual evocation of human behavior and capacities.

According to the European Commission's Communication on AI, a panel of high-level European experts in the field of AI:

"Artificial intelligence (AI) refers to systems that display intelligent behavior by analyzing their environment and taking actions – with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g., voice assistants, image analysis software, search engines, speech,

---

[2] For a more in depth understanding of McCarthy interpretation of the term "Artificial Intelligence" see the Journal Article *What is Artificial Intelligence?* by John McCarthy, November 2007.

and face recognition systems), or AI can be embedded in hardware devices (e.g., advanced robots, autonomous cars, drones, or Internet of Things applications)" (European Commission, 2018: Section 1).

The intelligence referred to when talking about AI is a composition of learned patterns that aims at replicating human intelligence for a specific purpose. Thanks to the continuous evolution and improvement of these systems, the imitation of human intelligence can reach a level in which it is applied to realize tasks, usually conducted by human beings, more efficiently and objectively, overcoming human limitations (Mika et al., 2019).

According to AlgorithmWatch, a non-profit advocacy organization dedicated to the field, the socio-economic impact of AI systems on individuals and societies is captured more accurately by automated and algorithmic decision-making systems (Spielkamp, 2019). Algorithmic-based automated systems, or Automated decision-making systems (ADM), are systems allowed to independently plan and decide processes to achieve a specified goal without micro-management, thanks to the way in which their algorithms are programmed. Automated decision-making refers to delegating decisional functions from a human agent capable of accountability to a non-biological form of information processing designed by humans according to specific rules of computation (Loi & Spielkamp, 2021). This delegation functions through automated means without human involvement, thanks to the use of data sets of factual or inferred data, which AI could provide. Automated Decision-Making systems (ADM) are, therefore, a socio-technological framework not limited to its technology but that encompasses a decision-making model on which it is based, an algorithm that converts the model into a computable code, the data introduced used as an input, and the political and economic circumstance in which it is applied and on which it has an impact (Spielkamp, 2019).

Considering its potentiality, ADM systems could be beneficial for people's life if all subordinate and cognitive tasks adequately contribute to the goal prefixed by the human agent that programmed them. Thanks to the chance it has to augment administrative efficiency, objectivity, and rationality while reducing bureaucratic times if used with appropriate management and control as a mere tool for humans (Loi & Spielkamp, 2021). Indeed, the independence of ADM systems is relative, given that

the algorithms upon which they are based are inherently relational and that they are created by human endeavor and operate according to human knowledge (Janssen & Kuk, 2016). Following this reasoning, these systems represent and shape the human limits and biases they pretend to overcome. Therefore, the use of ADM poses a series of social and political challenges that arise from the doubts about how transparent, open, accountable, and participatory a government in which ADMs are used as a standard tool for practice can be.

Relying on the collection of data and big data from individuals in society, ADMs started to be introduced in governmental practices to detect civilians' behaviors to assess whether they should receive welfare benefits, what kind of insurance they deserve, to calculate risks of fraud, or to evaluate the possibility of them committing acts of terrorism or criminal recidivism. Giving such responsibilities to ADMs raises social and ethical questions about how the data are attained and how reliable, unbiased, and non-discriminatory the decision taken with those data can be (European Commission, 2023). Already existing social biases and discriminatory patterns are reproduced, amplified, and integrated by such systems, as identified in various administrative practices in which ADMs are used.

The European Union promotes, in this sense, a human-centric approach to the development and use of AI and ADMs in which the systems should be lawful, robust, and in accordance with EU values and fundamental rights (European Commission, 2023).

3.2 Good Governance

The term governance is often used as a synonym for government, despite the two having significantly different meanings. The term government changed its significance with time, moving from representing the system by which something is governed in the mid-sixteenth century to the personification in the eighteenth century of the governing authority that leads a state, organization, or institution. On the other side, the term governance, which lost importance between the seventeenth and eighteenth centuries, reacquired significance in the 1980s with the introduction of globalization. Under these circumstances, the term governance, which refers to the process of

decision-making and implementation of those decisions, was now linked to the process of governing by sustainable development models (Iftimoaei, 2018).

Consequently, thanks to the interest demonstrated by international organizations (IO) and non-governmental organizations (NGOs) by the mid-1980s, the concept of Good Governance has been (re)introduced, reviving Aristotle's conceptualization of the notion, representing the process of governing in which adherence to the rule of law is emphasized. The term was officially introduced in public discourse in 1989 by World Bank President Barber Conable, that described it as a *"public service that is efficient, a judicial system that is reliable, and an administration that is accountable to its public"* (Sub-Saharan Africa, 1989: xxi). Since its beginning, the term good governance has concerned the economic and political spectrum. For this reason, it was of great interest to organizations such as the World Bank, the IMF, and the UNDP, being used as a reference for governance for developing countries.

To our interest, in this paper focus will be given to the political spectrum of good governance, which under a qualitative perspective, entails government architecture, democratization, and human rights protection (Reif, 2004). It compounds how public institutions should conduct public affairs respecting civilians' fundamental rights and democratic principles. According to the United Nations Office of the High Commissioner of Human Rights (OHCHR), the key attributes of good governance are transparency, responsibility, accountability, participation, and responsiveness (OHCHR, n.d.).

More extensively, a decision made through a decision-making process of good governance is participatory, consensus oriented, accountable, transparent, responsive, effective, efficient, equitable, inclusive, and legitimated by the rule of law. Good governance also enables decision-making processes, guaranteeing the respect of good governance principles not only in the final decision but in the whole process with which the decision is taken.

3.3 Ethics

In the landscape of Artificial Intelligence, it is intended, as Professor Adela Cortina Orts[3] underlines, that ethics should orient the application of technology and that the creation of a trustworthy AI (or good AI) depends on the complementation of technical progress and ethical progress (Cortis Orts, 2019).

For the purpose of this research, ethics will be intended as the five ethical principles that, according to the AI4People *Ethical Framework for a Good AI Society: Opportunities, Risks, and Recommendations*[4], should be a foundation for the development and adoption of good AI. These principles resemble the four core principles of bioethics, which according to Floridi, is the area of applied ethics that approximates most to digital ethics (Floridi, 2013). To address the concerns related to the use of AI, the principle are beneficence, non-maleficence, autonomy, justice, and explicability, with the latter as an addition to the normative principles of bioethics (Floridi et al., 2018).

The first ethical principle to consider is *beneficence,* which stands for the promotion of well-being, preservation of dignity, and the sustainment of the planet. In it, AI should benefit humanity and always represents the common good. Secondly*, non-maleficence* as an ethical principle attributed to AI represents the absence of harm or potential negative consequences to its use. The harm taken into consideration in this sense goes from the infringement of privacy to the repercussive self-improvement of AI. The avoidance of harm is understood as coming from the human being that programs the machine and the machine itself, both in cases of intentional and unpredicted harm.

A*utonomy*, corresponding to the third principle analyzed, indicates the equilibrium between the delegated decision-making power and the conserved one. According to this principle, the development of AI should allow and promote the autonomy of all human beings while controlling the autonomy of machines. In it, human beings should deter the power to choose the level of autonomy to be given to AI. Furthermore, the

---

[3] Prominent figure of the Spanish scholar scene in the field of ethics and moral philosophy, professor of Ethics and Legal, Moral and Political Philosophy at the University of Valencia since 1986.
[4] AI4People Framework, an Atomium-EISMD initiative, captures the meaning of 47 principles contained in six high-profile documents to lay the foundation for "Good AI Society".

ethical principle of *justice* invokes the importance of avoiding all kinds of discrimination and the equal share of benefits created by AI. It also entails the respect of all parties subject to the use of AI according to standards of solidarity, equality, and fairness.

Lastly, *explicability* implicates transparency, accountability, liability, and intelligibility principles. This principle complements those mentioned above, guaranteeing a level of responsibility for developing and deploying AI and assuring transparency apart from the systems' traceability, allowing attributing accountability in case of undesirable outcomes and for the ease of comprehension of affected subjects (Floridi et al., 2018).

4. RESEARCH DEVELOPMENT
4.1 Legislative overview of ADMs application in European public administrations

The digital transformation of governments is today a key objective for the European Union. In the past decades, new forms of technology reached all sectors of European society, including the public administrations of member states. While extremely fruitful in reducing expenditures and bureaucratic times, AI and specifically ADM systems in public administration, raised concerns over their responsiveness to the law and the respect for human rights and ethical principles core to the European Union.

In this scenario, the EU has some of the most extensive and valuable legislation in the world regarding data protection and digitalization, accompanied by various attempts to widely regulate the use of AI. In this respect, the most important and recognized form of regulation on the issue is the General Data Protection Regulation (GDPR), a widely accepted benchmark for data protection both in the public and private sectors, which applies to all twenty-seven countries of the European Union and countries of the European Economic Area (EEA).

Oriented towards the achievement of the Lisbon Strategy[5], the European interest in digitalization started in 2001 with the creation of the program e-Europe 2002[6], which aimed at the stimulation and divulgation of the use of the Internet, with attention to consumer protection and information. This was followed in 2010 by the creation of the *Digital Agenda* under the mandate of José Manuel Durão Barroso as President of the European Council. Durão Barroso incentivized the construction of a digital single market for the achievement of economic and social benefits with the creation of *Horizon 2020* (H2020)[7], a research and innovation funding program that operated between 2014 and 2020 with a budget of nearly 80 billion euros, of which 1.1billion was dedicated to research around AI (Guinea Bonillo, 2023).

While innovation and development for digitalization have been highly discussed since the start of the twenty-first century, it was only in 2018 that the interest veered to the future of AI for Europe (Guinea Bonillo, 2023). On the 25th of April 2018, the EU Commission shaped the Communication on *Artificial Intelligence for Europe*, delivered to all EU institutions in which the framing of AI, its potentiality, and the direction of its use was agreed upon (Artificial Intelligence for Europe, 2018).

However, when looking at AI's rapid expansion and development, more specific regulations that tackle the ethical impact these digital systems have on human beings seem lacking. This obliges us to refer to cardinal EU legislations, such as the Treaty of the European Union or its Charter of Fundamental Rights, to assure that the development of these technologies represents a common good for society. In this sense, Article 2 of the Treaty of the European Union stipulates that:

> *The Union is founded on the values of respect for human dignity, freedom, democracy, equality, the rule of law, and respect for human rights, including the rights of persons belonging to minorities. These values are common to the Member States in a society in which pluralism, non-discrimination, tolerance, justice, solidarity, and*

---

[5] *"To become the most competitive and dynamic knowledge-based economy in the world, capable of sustainable economic growth with more and better jobs and greater social cohesion"* (Presidency conclusions, Lisbon European Council, 23 and 24 March 2000). To which at the Gothenburg summit in June 2001 it was added *"and a sustainable environment"* (Directorate-General for Internal Policies et al., 2009).

[6] Followed then by *e-Europe 2005* approved by the Heads of State and Government at the Seville European Council in June 2002 as a new step in development of the information society.

[7] For more information refer to https://research-and-innovation.ec.europa.eu/funding/funding-opportunities/funding-programmes-and-open-calls/horizon-2020_en [Accessed last: 12/05/2023].

equality between women and men prevail (Consolidated Version of the Treaty on European Union, 2012).

Being the Treaty binding for all Member States implies that the prescriptions mentioned in the article must be observed collectively in all sectors, including in the use of AI. To underline this necessity, the preamble of the Charter of Fundamental Rights of the EU stresses the necessity to *"strengthen the protection of fundamental rights in the light of changes in society, social progress and scientific and technological developments"* (Charter of Fundamental Rights of the European Union, 2000)*.*

Using existing legislation that does not mention the use of AI to regulate algorithmic processes is necessary, but not ideal, to guarantee citizens' protection when subjected to the use of AI. This is the case, for example, of the protection against discrimination, inequalities, or the protection of citizens' right to good governance. Already existing legislation needs binding procedural rules to guarantee the endurance of such rights and the possibility of applying them to AI scenarios. As in the case of Article 41.1 of the Charter of Fundamental Rights of the EU concerning the right to good administration, which lacks binding procedural rules to guarantee citizens' right to good administration. Their absence makes the endurance of the aforementioned right impossible, which is highly problematic with the rise of digitalization in the public sector (European Parliament. Directorate General for Parliamentary Research Services, 2022).

Intending to approach the issue of AI more directly, in February 2020, the European Commission released the document: *White Paper on Artificial Intelligence - A European Approach to Excellence and Trust,* which in addition to previous texts, acknowledges the safety and liability aspects of AI. This document presents the High-Level Expert Group guidelines for the appropriate use of AI, which requires technical robustness and safety, human oversight, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental well-being, and accountability (European Commission, 2020).

However, these recommendations do not always align with EU priorities, as with the launch of the draft of *The Interoperable European Act* of November 2022, which had

the specific objective of linking data sources across Europe for use in public decision-making without, however, tackling upon how this distribution and exchange could ensure data quality and protect the subjects involved (Hofmann, 2023).

Additionally, according to Chapter 3 Article 22: *Automated individual decision-making, including profiling* of Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016, General Data Protection Regulation (GDPR):

> *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* (General Data Protection Regulation, 2016).

This impedes the processing of data without one's consent, except for the case of a contract between the data subject and the data controller, in case of given explicit consent, or if European or national legislations authorize the process, and if the rights and freedoms of the subject are protected in such processes.

While it is clear that technology advances at a much higher pace than inherent regulation, in April 2021, the European Commission proposed a body of laws named the *Artificial Intelligence Act* which aims at regulating AI with a proportionate horizontal regulatory approach that looks at regulating the use of AI. It aims to guarantee consistency with existing EU laws and guidelines and to ensure a high level of protection for fundamental rights by addressing various sources of risks through a risk-based approach (European Commission, 2021). This proposal is currently still under review and, therefore, yet to be enforced or binding. On Thursday, the 11th of May 2023, the Internal Market Committee and the Civil Liberties Committee voted in favor of the Act, adopting a final draft negotiating mandate, which is predicted to be voted upon in Parliament Plenary Session between the 12th and the 15th of June 2023 (AI Act, 2023).

4.2 Case studies

In this chapter, a series of case studies that illustrate the practical application of the concepts previously discussed in the thesis will be analyzed. Through these cases, there will be examined in detail the different facets of the topic of research and illustrate how the theories are put into practice in the real world. The selected cases were

carefully chosen for their relevance and particularity. They represent particular applications of ADM systems in different contexts of public administration, providing a unique opportunity to explore specific aspects of the research topic. Through the in-depth analysis of these cases, the study aims the enrichment of the academic debate, and it offers a practical perspective on the ethical impact of using ADMs in public administration for citizens.

I. The case of *SyRI*, ADMs for fraud detection in the Netherlands

In 2014 an amendment to the Work and Income Structure Implementing Organization Act (SUWI Act) was passed in the Dutch Parliament and the Senate. The amendment authorized the use of SyRI, acronymous for *Systeem Risico Indicatie* (System Risk Indicator), a legal instrument used by the Dutch government to detect various forms of social welfare fraud, including allowances, social benefits, and tax fraud (van Bekkum & Borgesius, 2021).

While Article 64.1 of the SUWI Act permitted designated government agencies to work together to prevent and combat the unlawful use of government funds and government provisions in the field of social security and income-related schemes and to prevent and combat tax fraud and the noncompliance with labor laws, Article 65 gave authority to do so by applying the SyRI system (*SUWI Act*, 2002). With the approval of such laws, the Dutch Data Protection Authority and the Council of State raised concerns about the extent of power and executive discretion for the government that the implementation of the law would entail.

As an ADM system applied to public administration, the SyRI system had immense information power, acquiring and cross-referencing data from national databases of different executive bodies and filling in dossiers on citizens without their consent or knowledge. Municipalities, the Employee Insurance Agency, the Social Insurance Bank, the Ministry of Social Affairs and Employment (SZW), the Immigration and Naturalization Service, and tax authorities had access to the system that could open an investigation on any specific citizen in case of suspected fraud. For the implementation of the system in a specific neighborhood, two agencies had to cooperate and ask for permission from the SZW with no obligation to inform subjected

citizens. SyRI's indications were then tested for false positives by the SZW before being investigated further on a case-by-case model by the agencies themselves (Jenny Gesley, 2020).

Given the high ethical risks posed by the use of the system, expressed in the obscurity it involved, that impeded citizens from knowing whether they were being investigated, and the exacerbation of discrimination and biases promoted by the exclusive application of the system in low-income neighborhoods (Jenny Gesley, 2020); the Netherlands Committee of Jurists for Human Rights (NJCM), together with other civil society organizations such as the Platform for the Protection of Civil Rights (Platform Bescherming Burgerrechten), and two natural persons jointly denounced the SyRI case. Thanks to public campaigns and the exposition of the case, in July 2018, the Federation of Trade Unions in the Netherlands (FNV) joined the coalition (Wisman & Hendrickx, 2020).

Rotterdam has been the first city to stop using SyRI at a local level in July 2019, calling the attention of representatives at the United Nations, such as Philip Alston, UN Special Rapporteur on Extreme Poverty and Human Rights, who wrote in the form of *Amicus Curiae* on the 26th of September 2019 a letter addressing the District Court of the Hague and bringing the case of SyRI all over the news (Alston, 2019).

This initiative led to the presentation, on the 5th of February 2020, of the NJCM c.s./De Staat der Netherlanden case against the Dutch government to establish a legal precedent for the protection of citizens from risk profiling and predictive policing practices (Wisman & Hendrickx, 2020). As the case was presented, the Court detected the blurriness of the system and its law, acknowledging the fact that the secrecy of the risk models made it impossible for citizens to defend themselves against the existence of a case against them, and ordered the immediate halt of the use of the system (AlgorithmWatch, 2020).

The Court appealed to general principles of the Charter of Fundamental Rights of the European Union and the GDPR. It explicitly determined SyRI to be in violation of Article 8.2 of the European Convention of Human Rights (ECHR), which establishes that:

> ARTICLE 8 Right to respect for private and family life *2. 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* (Convention for the Protection of Human Rights and Fundamental Freedoms, 1950).

The ECHR legislation regarded the application of SyRI as insufficiently transparent and verifiable, and therefore SyRI legislation as unlawful and without binding effect as in violation of higher law, not providing sufficient safeguards to protect citizens' privacy rights. In addition, the court declined to order the government to divulge the SyRi risk models used in specific SyRi projects, and the court did not order the destruction of the SyRi-collected data. The Netherlands announced on the 23rd of April 2020 that it had no plans to appeal the case's verdict (NJCM c.s./De Staat der Netherlanden, 2020).

After the verdict, Dutch public authorities announced that they would revise their fraud system, and many municipal councils started testing the legality of systems similar to SyRI (Wisman & Hendrickx, 2020). The rise of attention and concerns, both at the citizen and authority level, against risk profiling practices brought a significant public and political debate on the use of ADM systems in the country in public administration, serving as a precedent to consider for future applications.

II. The case of *Buona Scuola*, ADMs for the allocation of professional destinations in Italy

On the 13th of July 2015, the Italian government passed Law 107/2015, legislation that aimed at reaffirming the central role of schools, improving students' skills, countering socio-cultural inequalities, preventing dropouts, and guaranteeing the right to education. The implementation of this legislation, named *Buona Scuola* (Good School), implicated, according to Article 1.108 of the law, an extraordinary territorial and professional mobility plan on all vacant posts of common, support, and enhancement teaching managed by the Italian Ministry of Education (MIUR), at the time led by the Minister of Education Stefania Giannini (Legge 107/2015, 2015).

After an initial ordinary phase of recruitment, the reform plan was to be aimed, on the one hand, at filling uncovered common and support posts and, on the other hand, at filling additional posts intended for the enhancement of the educational offer and the coverage of temporary substitutions in primary and secondary schools through derogatory procedures. Unlike previous years, this mobility plan, which resulted in 210,000 transfer requests for the 2016/2017 school year, used an algorithmic-based system of ADM for the new allocation of teachers at a national level. The algorithm, whose name we disown, has been developed by the two companies Finmeccanica and Hp Enterprise Services Italia, winners of the MIUR fund for IT process management of 117 million euros (Zunino, 2019).

The ADM system was meant to process all necessary documentation of each applicant given by the MIUR and consider, following the law, the score achieved through merit and, subsequently, the list of fifteen preferences selected by each candidate at the time of the transfer application. However, the system resulted in working differently, prioritizing, in some cases, the personal predilection of candidates and bypassing the merit score achieved during the national examination or completely ignoring the preferences of the top-listed in others. Resulting in the detriment of the meritocratic criterion and the unjust displacement of innumerable teachers (De Angelis, 2019).

According to the examination done on the algorithm by a number of technicians at La Sapienza and Tor Vergata Universities, under the request of the Teachers Union, the ADM system was developed infringing the most basic parameters of programming and unnecessarily ostentatious, redundant, and unjustifiable. The engineers also observed the inconsistency in the programming of the system, manipulated in a fragmented manner by multiple programmers, which resulted in an unstable and chaotic scheme (Salvucci et al., 2017).

Thanks to the evaluation and after multiple complaints on the crypticity and unfairness of the algorithm by affected subjects, the case was presented first to the Regional Administrative Court (TAR) of Puglia and subsequently to TAR Lazio on the allegations of lack of transparency and bias to the meritocratic criterion of assignation (De Angelis, 2019), being then rejected (Sezione III bis, 10 Settembre 2018, n. 9227, 2018).

The appeal was then filed in front of the Council of State[8], in which the foundation of the claim laid in the violation of the principles of impartiality, transparency, and publicity, arising from the lack of knowledge of the mechanism adopted for the allocations, was confirmed (Mangano & Pignotti, 2022). In accordance, ruling No. 8472/2019 annulled the mobility procedure of the extraordinary recruitment plan under the program of *Buona Scuola* (Sentenza 13 Dicembre 2019, n.8472, 2019). This proceeding stressed the importance of implementing a higher level of digitalization in public administration to improve the quality of services delivered to citizens and users, both promoted at the national and European levels. In it, the Court individuates the use of algorithms systems in public administration useful to achieve greater speed, transparency, and efficiency, especially in the case of repetitive, serial, and discretion-free procedures. However, it recognized the application of this specific algorithm as unlawful, given the lack of transparency of its source code and use, and condemned its usage as contradicting the principle expressed in the GDPR and the requirements contained in Law 107/2015 to which it should refer (Sentenza 13 Dicembre 2019, n.8472, 2019).

The total computerization and human exclusion of MIUR's administrative procedure of professional mobility undoubtedly created multiple problems for the subjected teachers, who saw themselves obliged to follow the unjust allocation made by the system not to lose their position in national rankings. However, despite the critiques and the result of the sentence, the lesson has not been learned, and the MIUR, like other national administrative authorities, continues to implement ADM systems for the conduction of decision-making procedures, as in the case of the assignation of substitute teaching staff for the 2022/23 scholar year[9]. This system, like the one used in 2016, seems to be structurally problematic, chaotic, and obscure, and is the source of unemployment, incorrect distribution of roles, and the assignment of nonspecialized teachers to students with disabilities (MIUR Istruzione, 2022).

---

[8] Body of constitutional importance that retains both jurisdictional functions and advisory duties.
[9] It has been confirmed that the same algorithm will be used by MIUR in the academic year 2023/2024, with some modifications not yet fully disclosed (Talarico, 2023).

III.   The case of *BOSCO*, ADMs for the distribution of financial aid in Spain

In November 2017, the office of the Spanish Secretary of the State for Energy released BOSCO, an algorithmic-based software for decision-making made available to companies providing electricity to determine eligibility for the already existing *Bono Social*[10], a discount mechanism applied to the voluntary price for the small consumer (PVPC) with a maximum energy limit in the billing period. *Bono Social* aims the protection of vulnerable groups of consumers using the conditions established by the Royal Decree Law 897/2017 of the 6th of October, which regulates the status of vulnerable, severely vulnerable, and socially excluded consumers, to establish eligibility for this specific financial aid (Comisión Nacional de los Mercados y la Competencia, 2023).

The social bonus for electricity was defined in Royal Decree-Law 6/2009, of the 30th of April 2009. Its purpose was to financially favor specific electricity consumers covered by the Tariff of Last Resort and who met specific social, consumption, and purchasing power characteristics determined by the Administration. The not exclusive entry routes for the accreditation can be summarized in three main categories: to perceive low incomes, to be certified as a large family, or to be a beneficiary of a minimum disability or retirement pension with a minimum income.

This ADM system, based on the source code with which it was created, processes all the documentation needed to determine whether a candidate for the provision of such financial aid is eligible for it and expresses its final judgment by selecting those that, according to the system, are entitled to it without any sort of human intervention. Designed to accelerate and ease the entitlement process for public utility companies, it appears to commit errors in the assignment of aid, automatically discarding from the selection process cases that perfectly satisfied all requirements without giving explanations. These difficulties have been recognized thanks to the numerous complaints made through the ombudsman (Defensor del Pueblo, 2022).

---

[10] Introduced on the 1st of July 2009

The inherent and additional problem of the application of this system is that citizens applying for financial aid do not interact with public administration but relate directly with electrical companies, which do not have the competency or power to change the result of the algorithm or to know the reason behind determined choices.

The integration of BOSCO in the process of assignation of the *Bono Social* is undoubtedly one of the reasons that brought the remarkable decrease in the number of people to whom this aid has been granted. Indeed, before its integration, approximately 2.4 million people were beneficiaries of this benefit (CNMC blog, 2014). This number radically decreased to 1.2 million in 2022, according to the Spanish National Markets and Competition Commission (CNMC) (Comisión Nacional de los Mercados y la Competencia, 2022).

After multiple complaints, Civio, a Spanish non-profit investigative newsroom and citizen lobby, analyzed the system discovering its inefficiency based on the design of its algorithm. Its inefficiency was perpetrated by the denial of aid to eligible candidates. As a result, Civio asked the government BOSCO's source code to identify the reason behind those mistakes. The request passed through three different governmental ministries before arriving before the Spanish Committee for Transparency and Good Governance, which refused to share the code on the bases of Article 14.1J of the Law on Transparency and Access to Public Information[11], which establishes professional secrecy, intellectual and industrial property as one of the limits to the right of access to information protected by transparency (Civio, 2019).

Consequently, in July 2019, Civio filed an administrative complaint to the Central Contentious-Administrative Court, arguing that any ADM used in public administration should have a transparent and accessible source code by default, as it is for any legislation. The Court responded in 2022, rejecting the allegations, alleging that the release of the code, in addition to contravening intellectual property, would affect both public safety and national defense and sentenced the NGO to a fine, to which they appealed once again (Civio, 2022).

---

[11] Article 14. Ley 19/2013, of December 9: Limits to the right of access. *1. The right of access may be limited when accessing the information would be detrimental to (j) Professional secrecy and intellectual and industrial property* [translated from Spanish] (Transparencia de La Actividad Pública, 2013).

While the use of BOSCO is still ongoing, Civio and CNMC created an informative tool to put at the disposal of electricity consumers to check their eligibility for the *Bono Social* prior to the submission of data to BOSCO through the electricity companies and to suggest the best and more effective way to apply according to each case (Civio, 2018). The insistence from civil society organizations to bring this case to process aims to halt the use of this obstructive AI system to create a precedent and set the standards for future ADMs used in public governance in the country.

IV. The case of RADAR-iTE, ADMs for risk assessment and recognition of possible terroristic threats in Germany

As a result of the rise since 2012 of the number of militant Salafists recognized by the police as potentially violent, in February 2015, the German Federal Criminal Police Office Bundeskriminalamt (BKA), in collaboration with the Forensic Psychology Working Group of the University of Konstanz, launched a project for the creation of a risk assessment system to prevent terrorist attacks in the country (Bundeskriminalamt, 2017). The project brought to the creation of RADAR-iTE (Rule-based analysis of potentially destructive perpetrators to assess acute risk – Islamist terrorism), an algorithmic-based ADM system. Its first operational version was developed in September 2016 following an intense investigation on risk and protective factors for terrorist or comparable acts of violence, creating a user-oriented design, a test of practicability by criminal investigators, and examining the discriminatory power based on ROC analysis[12]. Following its first implementation at a national level in 2017, between the end of 2017 and the spring of 2019, the BKA and the Forensic Psychology Group of the University of Konstanz conducted a first evaluation of the system, as required by the German Ministry of Education and Research, that lead to the upgrade of RADAR-iTE to RADAR_iTE 2.0 (Sadowski et al., 2021).

The program, used by state and federal police officers and criminal investigators, requires a high level of standardization to ensure homogeneous assessments and the

---

[12] ROC Analysis as in Receiving and Operating Characteristic Analysis is based on a ROC curve, a graph relating the sensibility and specificity of a system at different level of cut-off.

comparability of results. This instrument is indeed able to follow a standardized calculus based on already processed information or additionally legally collect information regarding the subject's observable behaviors rather than personal characteristics such as religiosity or attitude to conduct the assessment of risk (Bundeskriminalamt, 2017). It comprises two elements, a case chronology, a structured compilation of information about a specific case ordered chronologically, and a risk assessment sheet for evaluating shifts. The case chronology must comply with objectiveness, completeness, and transparency for the precise and impartial description of the subject. At the same time, the risk assessment sheet provides information about specific thematic areas, such as violence against individuals, relation with authorities and institutions, and military affiliation, with relative time dimensions.

The 2.0 version of the system analyses both risk characteristics, which sum (+1), and protective characteristics[13], which detract (-1), to avoid overestimation of risk and stigmatization. The summed value of the number of +1 and -1 is then used to classify the subject in a specific range of risk. To the quantitative evaluation is added a qualitative evaluation that considers some specific characteristics or combination of characteristics as *red* or *green flags* (Sadowski et al., 2021).

In its reviewed version, the platform was also submitted to a content validity test to assess the relevance of the prioritization of individuals from the Islamist spectrum as potential committers of politically motivated attacks. Furthermore, a content validity in this respect, a parallel system called RISKANT[14] has been created as a framework that guarantees the case-by-base assessment of each classification made by RADAR-iTE previous to the application of any measure (Fernandez & de Lasala, 2021). The nationwide implementation of RADAR-iTE was also complemented by training programs for officials that underline the scope of the system as merely a prioritization system that must be accompanied by the intensive treatment of each case from human experts.

---

[13] Protective characteristics are those associated with law-abiding individuals (e.g., stable job, integration in local society, cooperation with the law)
[14] Risk analysis of those inclined to act on Islamist motivations (RISKANT).

Although the system has been demonstrated to be safe for citizens and in line with ethical principles, thanks to the extensive investigation and tests to which it has been subjected, its accuracy is not absolute and presents some limits. It is indeed core to the program's functionality that each investigated subject is already present in police archives resulting in the impossibility of detecting unknown potential risk operators. Furthermore, the algorithm of the program has been trained with a majority of male cases, given the limited female caseload. This results in the potential lack of detection of female cases (Sadowski et al., 2021).

As lines of error can always be processed, data can be insufficient and incorrect results can potentially appear. It is known that a case-by-case analysis must accompany the correct use of RADAR-iTE and that the planning and implementation of measures should always be determined by human beings and not the system itself. Meeting scientific criteria and being used following its original scope, RADAR-iTE has improved systematization and standardization, building a more efficient, objective, transparent, and comprehensible structure of prioritization of cases that eases the work of experts, fills the communication gaps between provincial bureaus, and that contributes to national safety (Sadowski et al., 2021).

4.3 Similarities and differences

The cases presented above are real examples of different applications of ADM systems in different areas of public administration. Each one represents the intent of European national governments to introduce digitalized systems in governance, applied with higher or lower levels of success. To assess the ethical impact these systems produce for citizens, the analysis will make use of the ethical framework created by the AI4People for "an AI good society" proposed above.

In the first instance, in the realm of *beneficence*, the ethical principle that established the necessity of a program to guarantee the well-being of citizens, it is seen that between the cases presented, both the cases of SyRI and the algorithms of Buona Scuola do not directly aim at the promotion of well-being, preservation of dignity or the sustainment of the planet, since their application is more related to organizational and

detective services. Conversely, the systems of BOSCO and RADAR-iTE aim to improve such features, one by relieving vulnerable citizens from paying electricity bills and the latter by preventing terrorist attacks and improving national security.

Even with the potentially good intentions of the systems, just one out of the four cases comply with the ethical principle of *non-maleficence.* Indeed, SyRI, the system of Buona Scuola and BOSCO all cause harm to citizens by infringing their privacy and depriving them of deserved or entitled benefits. Differently, the system of RADAR-iTE, being applied only to already filed individuals and not having definitive consequences, does not cause any harm to the German population. The difference that separates the German system from the other three cases is also crucial for assessing the systems' adherence to the third ethical principle of *autonomy.*

According to this principle, an ADM system should be balanced out by the presence of human beings in the entirety of decision-making processes in public administration. This balance is guaranteed only in the case of RADAR-iTE, in which, thanks to the continuous examinations conducted by human beings to test and improve the system and the implementation of the RISKANT process, are humans who determine the level of autonomy of the system itself. In the case of SyRI, the cross-over of data is not controlled by human beings, and no human intervention is present in the detection and signaling of potential fraud. In the algorithm of Buona Scuola, human intervention is wholly bypassed, both in the calculation and decision-making, as it is in the BOSCO system, which moreover distances itself from the public authority, being used directly by private companies.

For the correct application of ethical AI, systems of ADM should furthermore comply with the ethical principle of *justice*, which pretends the avoidance of discrimination. Regrettably, regarding this specific point, all the systems analyzed maintain some level of potential bias. Starting from the bluntest, RADAR-iTE does avoid discrimination, and its programming is tested for it. However, the scarcity of female data in police records did not allow an immersive training of the system in both the male and female spectrum, which could lead to a lower capacity of detection of potential female terrorists. Nevertheless, systems such as SyRI, BOSCO, and the algorithms of Buona Scuola present greater difficulties in complying with this ethical principle. The

appliance of SyRI only in specific low-income neighborhoods, BOSCO affecting vulnerable people's quality of life, and the algorithm of Buona Scuola being biased by teacher preferences rather than merit, constitute limits to the achievement of justice as equality, solidarity, and fairness.

Lastly, an ethical system should also be guided by the principle of *explicability*, it must be transparent, its source code and programming must be understandable and readable, and it must have a level of accountability for its actions. In this sense, RADAR-iTE can be characterized as a system fully complying with explicability, thanks to the openness of the way in which it works, the multiple investigations different experts underwent, and the fact that the implementation of measures is always left to authorities that become accountable for their decisions. Explicability, in this case, is achieved even if the system's source code has not been publicly published, thanks to the transparency with which the development of the system has been brought over. Differently, as underlined in the presentation of each of the remaining cases, explicability is an essential element determining their failure. The opacity of the processes that gave rise to the programs themselves, the obscurity of their operations, and the lack of accountability for their actions represent, indeed, the main problem of those systems. In the case of SyRI and Buona Scuola system, it has been the reason for which their use has been prohibited.

In the assessment of the adherence to ethical principles of the systems taken into analysis, it is possible to identify RADAR-iTE as the only system that could potentially bring benefit from its use and be deployed under the doctrine of good governance. This is permissible thanks to its limitations in scope, which prevents the extension of its use and gives more importance and power to authorities, easing their work without exceeding it. Differently, the other three forms of ADM do not comply with the values of ethics that enable good governance. While two of them are already prohibited, the BOSCO system, as well as other ADM systems used in Europe, is still in use and far from being abolished, which determines the importance of this assessment and the urgency with which actions in merit must be taken.

5. CONCLUSION

Throughout this paper, it is analyzed the application of ADM systems in European public administration and its ethical impact on civilians. Thanks to the bibliography examined, it has been possible to examine the different conceptions of the topic of ethics and principles of good governance, which set the standards for the correct application of decision-making processes under a liberal and democratic standpoint that put the human being at the center and prioritize the well-being of civilians. When these principles are applied to the use of AI in the public sector, the challenges these systems inherently encompass automatically arise.

When discussing the impregnation of technology in both public and private sectors and its inevitable and continuous development, it is possible to discover the importance of this investigation to bring to light issues that are to the public obscure and ambiguous. These issues, however, directly interest society, which is continuously subject to technological advances and implementation. In an era of New Digital Governance in which technology is a core element that determines the social order, how processes of governance are made, and how everything evolves and develops, assessing if the algorithmic systems used are ethical and in accordance with democratic values, is fundamental to assure the maintenance of the status quo.

According to the ethical principle of liberal utilitarianism, as seen above, the sole purpose for which power is rightfully exercised over society is preventing harm to others generated by one's action or inaction. This theory, which focuses on the consequences of actions, is perfectly suitable for studying the ethicality of AI systems, which by their actions or inactions, create consequences that affect governed populations. When this attention to consequences is applied to the modern era of New Digital Governance, it is possible to observe how technological advances in governance processes raise questions and doubts about the possibility and the inherent consequences of a fully automatized domination.

The consequences of the application of AI to public administration can result to be both positive and negative, depending on the prospect from which this application is observed and how a specific system is developed and then used. From the

perspective analyzed through this paper, the challenges created by this implementation concern the ethical spectrum and can be resumed in the non-compliance of systems with the ethical principles of beneficence, non-maleficence, autonomy, justice, and explicability. This non-compliance can lead to defects to the well-being of citizens, the infliction of direct harm, the technological overthrow of roles and positions reserved for human beings, the conception of injustices, discrimination, and biases, the violation of privacy, the rise of data determinisms, and so forth.

However, the primary aim of the instruction of AI or ADM systems in public administration is to increase the efficiency, objectivity, and rationality of existing processes to improve them and ease the work of public officials. It is in assessing whether these three factors are value-added to processes thanks to the implantation of technologies and whether this implementation can coexist with respect to human dignity and the preservation of civilian well-being. It is indeed essential to look towards the coexistence of these properties, as the application of AI systems in public administration should not solely serve the needs and desires of public authorities and officials. They should work in respect of human rights and to enhance the lives of citizens, not to hinder but to benefit them and to promote public interests.

However, this coexistence and the prioritization of public interest are not always easily guaranteed. The great speed with which technologies develop and mutate, acquire capacities, and become more independent does not coincide with the timeline with which this development is acknowledged, comprehended, and regulated. The lack of ad hoc regulations that normalize the use of artificial intelligence systems protecting citizens is an issue commonly shared worldwide. In this context, the European Union can vaunt multiple attempts to regulate not only the development of digital transformation but also to ensure the respect of human rights in them. Nevertheless, additional efforts and concrete legislation that prevent the misuse of AI systems are increasingly fundamental.

The conception of new legislations on the topic must be predictive and open to the evolution and new forms of technology that will eventually become part of our society. Limiting guidelines to correcting damages already created will always result in legal loopholes that do not permit the proper protection and promotion of civilians' well-

being. Especially when applied to public administration, technology should be entirely regulated and in accordance with already existing legislation. This importance comes from the fact that its application, implemented by public authorities, directly affects civilians and, in some cases, vulnerable groups, behaving as actual directives.

Given the advanced position of the European Union on the matter and the steps forward, with the creation of specific legislation and its enforcement, the Union could become a referent in the ethical use of AI systems that put human beings at the center. Despite the popularity of this hypothesis, which sees the EU as a trendsetter in the field, as it has been with the creation of the GDPR, additional efforts must be applied to guarantee the approval of comprehensive laws at the European and national levels that are then respected in both the private and the public sectors. Furthermore, the development and implementation of such laws should be closer to the timelines of technology, ensuring the maintenance of relevance.

From the current standpoint, the lack of specific regulations and the generality of existing rules lead to poor observance of ethical values in the development of systems that are then put at the service of society. In the European Union, civil societies and NGOs, at a national or international level, play a fundamental role in fighting for the respect of human rights and the preservation of human dignity. It is wrong that despite specific systems' clear unsuitability, obscurity, or their apparent faults, their hold is not immediate. However, it depends on the efforts, determination, and to a certain extent, the fortune of organizations that bring to court the determined system, as it is possible to observe in the cases presented.

To remedy this issue, each country should be much more attentive to the programs that its government or officials of public administration use. They should run tests and correct anomalies prior to the release of each system and maintain them under control, correcting possible problems that can arise at later stages. This line of work could perhaps lead to better management of AI that aims at accompanying human beings without replacing them. This has been demonstrated as possible in different applications, as in the example of the development of the German RADAR-iTE system, which is studied and tested not to harm citizens.

It is possible that technology is not yet ready to take the role of human beings in such positions as governance or public administration. However, it is certainly the case that the public sector is not ready to completely digitize its procedures because it cannot protect its citizens from the use of AI and the consequences it produces. In this respect, three main proposals could fill in the existing gaps and promote an appliance of AI systems in public administration that benefit both governments and citizens.

First, the extended promotion of transparency and accountability. Public authorities should establish a set of guidelines that require public disclosure of AI systems they deploy, bringing to light their purpose, the source code with which they are created, and the process with which they would function. This transparency should also allow internal and external auditability of such systems to guarantee their efficiency and correct functioning. Furthermore, to guarantee accountability, National Transparency Councils should be given more responsibilities and power in determining the compatibility of a system with ethical values. They should be counted as a reference to direct to in case of detected system misuse, and they should have the power to address the issue directly. Their work should, however, always be citizen-oriented, and in it, the public interest must always prevail.

Second, the stress of data protection. Public administration bodies should employ robust data governance frameworks to regulate the collection, storage, and use of personal data in AI systems. These frameworks should adhere to applicable privacy laws, promote data anonymization whenever feasible, and establish mechanisms for obtaining individuals' explicit and informed consent. In this regard, they should prohibit the process of fully automated data determinism for predictive policing, a proposal advanced in the under-revision AI Act. Furthermore, data sharing and interoperability between different agencies or departments, promoted by European digitalization plans, should be highly reviewed and integrated with regulations that protect citizens' privacy and prevent over-fixation.

Third, for the accurate implementation of previous proposals, it is necessary to generate a data culture in which citizens are informed about the data they share and about their capacity to decide what information they disclose. It is also fundamental to gear the digital gaps in society and promote digital inclusion with a holistic approach.

This promotion should guarantee equitable access to AI technologies and assure accessibility and affordability to technology, together with digital literacy. In this respect, the induction of digital education should aim at citizens' comprehension and knowledge about the applicability of ADM systems or other AI systems to public administration. Allowing the perception and analysis of outcomes would make the population more demanding of processes that follow ethical values and respect human well-being.

Only by taking into consideration these proposals authorities and citizens will be able to take advantage of the development and capacities of new technologies, channeling them to satisfy their needs and desires. While the complete exclusion of AI from public governance is not feasible or ideal, it is essential to recognize the need to take various steps to ensure its application is consistent with the principles of good governance. The significance of this analysis cannot be overstated, as the topic remains highly current, and the implications for the future are uncertain. More than passive observation is required for proactively navigating this dynamic landscape. We must actively engage in shaping and demanding the ethical and effective implementation of AI in public administration.

6. BIBLIOGRAPHY

AlgorithmWatch. (2020). *How Dutch activists got an invasive fraud detection algorithm banned*. AlgorithmWatch. Accessible at: https://algorithmwatch.org/en/syri-netherlands-algorithm/ [Accessed last: 15/05/2023]

Alston, P. (2019, September 26). *Brief by the United Nations Special Rapporteur on extreme poverty and human rights as Amicus Curiae in the case of NJCM c.s./De Staat der Nederlanden (SyRI) before the District Court of The Hague (case number: C/09/550982/ HA ZA 18/388)*. Accessible at: https://www.ohchr.org/sites/default/files/Documents/Issues/Poverty/Amicusfin alversionsigned.pdf [Accessed last: 15/05/2023]

Andrikopoulos, V. P., & Ifanti, A. A. (2020). New Public Management and Governance: Quo Vadis? *Journal of Public Administration and Governance*, *10*(3), 430. https://doi.org/10.5296/jpag.v10i3.17494

Basak, M. K., & Chakroborty, R. (2018). *Good governance: Some ethical issues.* Progressive Publishers.

Bentham, J., Burns, J. H., & Hart, H. L. A. (1996). *An introduction to the principles of morals and legislation.* Clarendon Press ; Oxford University Press.

Bianchini, D., & Avila, I. (2014). Smart Cities and Their Smart Decisions: Ethical Considerations. *IEEE Technology and Society Magazine, 33*(1), 34–40. https://doi.org/10.1109/MTS.2014.2301854

Bundeskriminalamt. (2017, February 2). *BKA Press release: New tool for risk assessment of potential violent criminals. RADAR-iTE (rule-based analysis of potentially destructive perpetrators to assess the acute risk—Islamist terrorism)* [Official webpage]. Bundeskriminalamt. Accessible at: https://www.bka.de/DE/Presse/Listenseite_Pressemitteilungen/2017/Presse2017/170202_Radar.html [Accessed last: 15/05/2023]

Charter of Fundamental Rights of the European Union, 364/01 (2000). Accessible at: https://www.europarl.europa.eu/charter/pdf/text_en.pdf [Accessed last: 15/05/2023]

Civio. (2018, June 27). *Una aplicación de Civio y la CNMC ayuda a los consumidores de electricidad a comprobar si tienen derecho al bono social y a solicitarlo.* Civio. Accessible at: https://civio.es/novedades/2018/06/27/una-aplicacion-de-civio-y-la-cnmc-ayuda-a-los-consumidores-de-electricidad-a-comprobar-si-tienen-derecho-al-bono-social-y-a-solicitarlo/ [Accessed last: 15/05/2023]

Civio. (2019, July 2). *Que se nos regule mediante código fuente o algoritmos secretos es algo que jamás debe permitirse en un Estado social, democrático y de Derecho.* Civio. Accessible at: https://civio.es/novedades/2019/07/02/que-se-nos-regule-mediante-codigo-fuente-o-algoritmos-secretos-es-algo-que-jamas-debe-permitirse-en-un-estado-social-democratico-y-de-derecho/ [Accessed last: 15/05/2023]

Civio. (2022, February 10). *La Justicia impide la apertura del código fuente de la aplicación que concede el bono social.* Civio. Accessible at: https://civio.es/novedades/2022/02/10/la-justicia-impide-la-apertura-del-codigo-fuente-de-la-aplicacion-que-concede-el-bono-social/ [Accessed last: 15/05/2023]

CNMC blog. (2014, October 10). *¿Cuántos usuarios se benefician del bono social de electricidad en España?* CNMC Blog. Accessible at:

https://blog.cnmc.es/2014/10/10/cuantos-usuarios-se-benefician-del-bono-social-de-electricidad-en-espana/ [Accessed last: 15/05/2023]

Comisión Nacional de los Mercados y la Competencia. (2022, May). *CNMC Data* [Estadística bono social]. CNMC Data. Accessible at: https://data.cnmc.es/energia/energia-electrica/bono-social/estadistica-bono-social [Accessed last: 15/05/2023]

Comisión Nacional de los Mercados y la Competencia. (2023, March 21). *Bono social eléctrico | CNMC* [Official webpage]. Comisión Nacional de Los Mercados y La Competencia 2022. Accessible at: https://www.cnmc.es/bono-social#que-es-el-bono [Accessed last: 15/05/2023]

Communication From The Commission To The European Parliament, The European Council, The Council, The European Economic And Social Committee And The Committee Of The Regions Artificial Intelligence For Europe, COM(2018) 237 final (2018). Accessible at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A237%3AFIN [Accessed last: 15/05/2023]

Consolidated version of the Treaty on European Union, C 326/15 (2012). Accessible at: https://eur-lex.europa.eu/resource.html?uri=cellar:2bf140bf-a3f8-4ab2-b506-fd71826e6da6.0023.02/DOC_1&format=PDF [Accessed last: 15/05/2023]

Convention for the Protection of Human Rights and Fundamental Freedoms, Rome, 4.XI.1950 (1950). Accessible at: https://www.echr.coe.int/Documents/Convention_ENG.pdf [Accessed last: 15/05/2023]

Cortis Orts, A. (2019). Ética de la Inteligencia Artificial. *Anales de la Real Academia de Ciencias Morales y Políticas*, *96*, 379–394.

De Angelis, M. (2019). Algoritmi nei concorsi pubblici: Il caso dei docenti che fa "scuola." *IUS in Intinere*. Accessible at: https://www.iusinitinere.it/algoritmi-nei-concorsi-pubblici-il-caso-dei-docenti-che-fa-scuola-23299 [Accessed last: 15/05/2023]

Dean, M. (2010). *Governmentality: Power and rule in modern society* (2nd ed). SAGE.

Defensor del Pueblo. (2022). *Defensor del Pueblo- Informe anual 2022. Volumen I* (M-8039–2023; Informe anual, pp. 99–103). Accessible at:

https://www.defensordelpueblo.es/wp-content/uploads/2023/03/Defensor-del-Pueblo-Informe-anual-2022.pdf [Accessed last: 15/05/2023]

Dent, C. (2020). Identity, Technology and their Confluence: Governmentality in the Digital Age. *Law, Technology and Humans*, *2*(2), 81–96. https://doi.org/10.5204/lthj.v2i2.1437

Directorate-General for Internal Policies, Committee on Employment and Social Affairs, & The Secretariat. (2009). *Briefing note for the meeting of the EMPL Committee 5 October 2009 regarding the exchange of views on the Lisbon Strategy and the EU cooperation in the field of social inclusion*. European Parliament. Accessible at: https://www.europarl.europa.eu/meetdocs/2009_2014/documents/empl/dv/lisbonstrategybn_/lisbonstrategybn_en.pdf [Accessed last: 15/05/2023]

Dunleavy, P., Margetts, H., Bastow, S., & Tinkler, J. (2005). New Public Management Is Dead—Long Live Digital-Era Governance. *Journal of Public Administration Research and Theory*, *16*(3), 467–494. https://doi.org/10.1093/jopart/mui057

European Commission. (2020). *White Paper On Artificial Intelligence—A European approach to excellence and trust*. Accessible at: https://commission.europa.eu/system/files/2020-02/commission-white-paper-artificial-intelligence-feb2020_en.pdf [Accessed last: 15/05/2023]

European Commission. (2021). *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules On Artificial Intelligence (Artificial Intelligence Act) And Amending Certain Union Legislative Acts*. European Commission. Accessible at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A52021PC0206 [Accessed last: 15/05/2023]

European Commission. (2023, February 7). *Automated decision-making impacting society* [Official webpage]. Competence Center on Foresight. Accessible at: https://knowledge4policy.ec.europa.eu/foresight/automated-decision-making-impacting-society_en [Accessed last: 15/05/2023]

European Commission. Directorate General for Employment, Social Affairs and Inclusion. (2017). *Quality of public administration: A toolbox for practitioners. Principles and values of good governance.* Publications Office. Accessible at: https://data.europa.eu/doi/10.2767/593135 [Accessed last: 15/05/2023]

European Parliament (2023) *AI Act: A step closer to the first rules on Artificial Intelligence | News | European Parliament*. (2023, May 11). [Press Released].

European Parliament News. Accessible at: https://www.europarl.europa.eu/news/en/press-room/20230505IPR84904/ai-act-a-step-closer-to-the-first-rules-on-artificial-intelligence [Accessed last: 15/05/2023]

European Parliament. Directorate General for Parliamentary Research Services. (2022). *Digitalisation and administrative law: European added value assessment.* Publications Office. Accessible at: https://data.europa.eu/doi/10.2861/643042 [Accessed last: 15/05/2023]

Fernandez, C., & de Lasala, F. (2021). *Risk Assessment in Prison*. European Union. Accessible at: https://home-affairs.ec.europa.eu/system/files/2021-04/ran_cons_overv_pap_risk_assessment_in_prison_20210210_en.pdf [Accessed last: 15/05/2023]

Floridi, L. (2013). *The Ethics of Information*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199641321.001.0001

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations. *Minds and Machines*, *28*(4), 689–707. https://doi.org/10.1007/s11023-018-9482-5

Guinea Bonillo, J. (2023). La Unión Europea y el desafío de la Inteligencia Artificial. In *Inteligencia Artificial y Derecho reflexiones jurídicas para el debate sobre su desarrollo y aplicación* (Vol. 1–41–56). DYKINSON, S.L.

Häyry, M. (1994). *Liberal utilitarianism and applied ethics*. Routledge.

Hofmann, H. C. H. (2023). Comparative Law of Public Automated Decision-Making. An Outline. *CERIDAP*, *1/2023*. https://doi.org/10.13130/2723-9195/2023-1-3

Hood, C. (1995). The "new public management" in the 1980s: Variations on a theme. *Accounting, Organizations and Society*, *20*(2–3), 93–109. https://doi.org/10.1016/0361-3682(93)E0001-W

Iftimoaei, C. (2018). *Toward a Sociological Approach of Good Governance*.

Janssen, M., & Kuk, G. (2016). The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly*, *33*(3), 371–377. https://doi.org/10.1016/j.giq.2016.08.011

Jenny Gesley. (2020, March 13). *Netherlands: Court Prohibits Government's Use of AI Software to Detect Welfare Fraud* [Web page]. Library of Congress,

Washington, D.C. 20540 USA. Accessible at: https://www.loc.gov/item/global-legal-monitor/2020-03-13/netherlands-court-prohibits-governments-use-of-ai-software-to-detect-welfare-fraud/ [Accessed last: 15/05/2023]

Kearns, M., & Roth, A. (2020). *The ethical algorithm: The science of socially aware alogrithm design*. Oxford University Press.

Kellis, D. S., & Ran, B. (2013). Modern leadership principles for public administration: Time to move forward: Modern leadership principles. *Journal of Public Affairs*, *13*(1), 130–141. https://doi.org/10.1002/pa.1453

Kitchin, R. (2016). The ethics of smart cities and urban science. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2083), 20160115. https://doi.org/10.1098/rsta.2016.0115

Larry, A., & Moore, M. (2021). Deontological Ethics. In *The Stanford Encyclopedia of Philosophy* (Winter 2021). Metaphysics Research Lab, Stanford University. Accessible at: https://plato.stanford.edu/archives/win2021/entries/ethics-deontological [Accessed last: 15/05/2023]

Lassalle, J. M. (2019). *Ciberleviatán: El colapso de la civilización democrática y liberal frente a la revolución digital* (Primera edición). Arpa.

Ley 19/2013 de transparencia, acceso a la información pública y buen gobierno, no. BOE-A-2013-12887, Jefatura de Estado (2013). Accessible at: https://www.boe.es/buscar/doc.php?id=BOE-A-2013-12887 [Accessed last: 15/05/2023]

Loi, M., & Spielkamp, M. (2021). Towards Accountability in the Use of Artificial Intelligence for Public Administrations. *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*, 757–766. https://doi.org/10.1145/3461702.3462631

Lynch, T. D. (2004). *Virtue Ethics, Public Administration, and Telos. 5*(4), 32–49.

Mangano, G., & Pignotti, S. (2022). Focus sentenze G.A. su decisioni algoritmiche – Digitalizzazione 4.0: Illegittimità del ricorso all'algoritmo quando la regola giuridica sottesa è incomprensibile. *IRPA*. Accessible at: https://www.irpa.eu/focus-sentenze-g-a-su-decisioni-algoritmiche-digitalizzazione-4-0-illegittimita-del-ricorso-allalgoritmo-quando-la-regola-giuridica-sottesa-e-incomprensibile/ [Accessed last: 15/05/2023]

Mantzouranis, K. (2018). Thucydides' Assessments of Pericles and Alcibiades as a Lesson in Leadership Ethics. *Polis: The Journal for Ancient Greek Political Thought*, *35*(2), 523–547. https://doi.org/10.1163/20512996-12340178

McCarthy, J. (2007). What Is Artificial Intelligence? *Computer Science Department*. Standford University. Accessible at: https://www.diochnos.com/about/McCarthyWhatisAI.pdf [Accessed last: 15/05/2023]

Mika, N., Nadezhda, G., Jaana, L., & Raija, K. (2019). *Ethical AI for the Governance of the Society: Challenges and Opportunities*.

MIUR Istruzione. (2022, September 21). *Supplenze Docenti GPS 2022-2023, errori algoritmo e ricorsi, Le Novità—MIUR Istruzione*. MIUR Istruzione. Accessible at: https://www.miuristruzione.it/28234-supplenze-docenti-gps-2022-2023-errori-algoritmo-e-ricorsi-le-novita/ [Accessed last: 15/05/2023]

Moor, J. (2006). The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years. *AI Magazine*, *27*(4), Article 4. https://doi.org/10.1609/aimag.v27i4.1911

Muellerleile Chris & Robertson Susan L. (2018). Digital Weberianism: Bureaucracy, Information, and the Techno-rationality of Neoliberal Capitalism. *Indiana Journal of Global Legal Studies*, *25*(1), 187. https://doi.org/10.2979/indjglolegstu.25.1.0187

NJCM c.s./De Staat der Netherlanden, C/09/550982/HA ZA 18/388 (District Court of Hague February 5, 2020). Accessible at: https://www.escr-net.org/sites/default/files/caselaw/ecli_nl_rbdha_2020_1878.pdf [Accessed last: 15/05/2023]

OHCHR. (n.d.). *About good governance* [Official webpage]. United Nations. Accessible at: https://www.ohchr.org/en/good-governance/about-good-governance [Accessed last: 15/05/2023]

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) (Text with EEA relevance), L 119 (2016).

Reif, L. C. (2004). *The ombudsman, good governance, and the international human rights system*. Martinus Nijhoff Publishers.

Riforma del sistema nazionale di istruzione e formazione e delega per il riordino delle disposizioni legislative vigenti., 15G00122 Camera dei deputati e Senato (2015). Accessible at: https://www.gazzettaufficiale.it/eli/id/2015/07/15/15G00122/sg [Accessed last: 15/05/2023]

Sadowski, F., Meier, H., Sonka, C., Witt, R., & Malzacher, J. (2021). RADAR-iTE 2.0: An Instrument of the German State Protection – Structure, Development and Current Stage of Evaluation. In R. Corrado, G. Wössner, & A. Merari (Eds.), *NATO Science for Peace and Security Series – E: Human and Societal Dynamics*. IOS Press. https://doi.org/10.3233/NHSDP210013

Salvucci, A., Giorgi, M., Barchesi, E., & Scafidi, M. (2017). *Perizia tecnica preliminare sull'analisi dell'algoritmo che gestisce il software della mobilità docenti per l'a.s. 2016/2017*. Accessible at: https://www.gildavenezia.it/perizia-tecnica-preliminare-sullanalisi-dellalgoritmo-che-gestisce-il-software-della-mobilita-docenti-per-la-s-20162017/ [Accessed last: 15/05/2023]

Sentenza 13 Dicembre 2019, n.8472, No. 8472 (Consiglio di Stato December 13, 2019). Accessible at: https://www.digies.unirc.it/documentazione/materiale_didattico/697_2020_1379_37385.pdf [Accessed last: 15/05/2023]

Sezione III bis, 10 settembre 2018, n. 9227, No. 9227 (TAR Lazio September 10, 2018). Accessible at: https://www.giurdanella.it/wp-content/uploads/2018/10/Tar-Lazio-Sez.-III-bis-10-settembre-2018-n.-9227.pdf [Accessed last: 15/05/2023]

Spielkamp, M. (2019). *Automating Society Report 2019* (Automating Society Report). AlgorithmWatch & Bertelsmann Stiftung. Accessible at: https://algorithmwatch.org/en/wp-content/uploads/2019/02/Automating_Society_Report_2019.pdf [Accessed last: 15/05/2023]

Sub-Saharan Africa: From crisis to sustainable growth: along-term perspective study. (1989). World Bank. Accessible at: https://documents1.worldbank.org/curated/en/498241468742846138/pdf/multi0page.pdf [Accessed last: 15/05/2023]

Talarico, R. (2023, February 16). *Algoritmo supplenze 2023*. La Scuola Oggi. Accessible at: https://www.lascuolaoggi.it/gps-graduatorie-provinciali-supplenze/algoritmo-supplenze-2023/ [Accessed last: 15/05/2023]

Tan, E., & Crompvoets, J. (2022). A new era of digital governance. In E. Tan & J. Crompvoets (Eds.), *The new digital era governance* (pp. 13–49). Wageningen Academic Publishers. https://doi.org/10.3920/978-90-8686-930-5_1

Thucydides. (2003). *The History Of The Peloponnesian War*. Accessible at: https://www.gutenberg.org/files/7142/7142-h/7142-h.htm#link2HCH0010 [Accessed last: 15/05/2023]

UN DESA. (2022). *United Nations E-Government Survey 2022: The future of digital government.* United Nations. Accessible at: https://desapublications.un.org/sites/default/files/publications/2022-09/Web%20version%20E-Government%202022.pdf [Accessed last: 15/05/2023]

Van Bekkum, M., & Borgesius, F. Z. (2021). Digital welfare fraud detection and the Dutch SyRI judgment. *European Journal of Social Security*, *23*(4), 323–340. https://doi.org/10.1177/13882627211031257

Wet structuur uitvoeringsorganisatie werk en inkomen [Structure of Implementation Organizations for Work and Income Act], Staatsblad (Stb.) (2002). Accessible at: https://wetten.overheid.nl/BWBR0013060/2020-01-01 [Accessed last: 15/05/2023]

Wisman, T., & Hendrickx, M. (2020, April 23). The SyRI Victory: Holding Profiling Practices to Account. *Digital Freedom Fund*. Accessible at: https://digitalfreedomfund.org/the-syri-victory-holding-government-profiling-to-account/ [Accessed last: 15/05/2023]

Zunino, C. (2019, September 17). Scuola, trasferimenti di 10mila docenti lontano da casa. Il Tar: "L'algoritmo impazzito fu contro la Costituzione." *la Repubblica*. Accessible at: https://www.repubblica.it/cronaca/2019/09/17/news/scuola_trasferimenti_di_10mila_docenti_lontano_da_casa_il_tar_l_algoritmo_impazzito_fu_contro_la_costituzione_-236215790/ [Accessed last: 15/05/2023]