

Article

Artificial Intelligence Risks and Challenges in the Spanish Public Administration: An Exploratory Analysis through Expert Judgements

Itziar Sobrino-García 

Department of Public Law, Facultad de Ciencias Jurídicas y del Trabajo, Campus Lagoas-Marcosende, University of Vigo, 36310 Vigo, Spain; isobrino@uvigo.es

Abstract: The expanding use of artificial intelligence (AI) in public administration is generating numerous opportunities for governments. Current Spanish regulations have established electronic administration and support the expansion and implementation of this new technology, but they may not be adapted to the legal needs caused by AI. Consequently, this research aims to identify the risks associated with AI uses in Spanish public administration and if the legal mechanisms can solve them. We answer these questions by employing a qualitative research approach, conducting semi-structured interviews with several experts in the matter. Despite the benefits that this technology may involve, throughout this research we can confirm that the use of artificial intelligence can generate several problems such as opacity, legal uncertainty, biases, or breaches of personal data protection. The mechanisms already provided by Spanish law are not enough to avoid these risks as they have not been designed to face the use of artificial intelligence in public administration. In addition, a homogeneous legal definition of AI needs to be established.

Keywords: artificial intelligence; public sector; digital transformation; Spain; public administration



Citation: Sobrino-García, Itziar. 2021. Artificial Intelligence Risks and Challenges in the Spanish Public Administration: An Exploratory Analysis through Expert Judgements. *Administrative Sciences* 11: 102. <https://doi.org/10.3390/admsci11030102>

Received: 8 July 2021

Accepted: 8 September 2021

Published: 16 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The public sector is involved in a process of digital transformation to adapt to digital services and the new demands of citizens. Governments are changing their procedures to improve service delivery, be more effective and efficient in their public policies, and also increase transparency, interoperability, and citizen trust (Mergel et al. 2019). Therefore, technological advances have caused public administrations to embrace a more dynamic and flexible position, and consequently, they need to adopt new innovative administration models, which can improve the reaction capacity of the public sector (Campos Acuña 2019).

However, public administrations have had problems adapting to technological advances, since in several cases there has not been a good political situation, nor have there been adequate organisational structures (Dawes 2013). In addition, there has not been a concern to promote smart governance (Criado 2021), understood as a new governance management paradigm based on the promotion of new technologies, practices, policies, and more efficient resources. The COVID-19 crisis has highlighted the value of digital administration, which has allowed European countries to improve their offer of digital services this last year, because of mobility restrictions and measures to prevent the pandemic.

In fact, the European Parliament has already described various technologies as key strategies of the 21st century to improve the digitisation process, among which is artificial intelligence (AI) (European Parliament 2019b), generally referred to a combination of machine learning techniques used for searching and analysing large volumes of data, algorithms, and automated decision-making systems (ADMS) able to predict human and machine behaviour and to make decisions (European Parliamentary Research Service 2019). Thus, AI has the potential to transform different aspects of governments, including interactions with citizens, the provision of services, the design of public policies, and

decision-making (Sun and Medaglia 2019; Valle-Cruz et al. 2020). For this reason, public authorities are introducing the use of these methods in various stages of public management. Different AI projects or experiences have been developed but in specific sectors, such as the field of health (Mesko 2017), sustainability (Nishant et al. 2020), labour market (Allam and Dhunny 2019), or mobility (Nikitas et al. 2020), which have generated a lack of global policy and synergies that help to solve complicated problems in the public sector.

Nevertheless, uncertainty about the potential impact of new technology and an existing legal framework not adapted for the new socio-technical scenario has generated a series of difficulties regarding the use of AI. At the European level, a legal answer has been explored to help solve the risks of AI, following a series of principles centred on the human being (Piñar Mañas 2019; Cerrillo i Martínez 2020b). Therefore, at the end of April 2021, the Commission presented a proposal for a regulation on AI, "Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and certain Union legislative acts". Due to the Commission being aware of the multiple applications that the use of AI can imply to users, it is vitally important for the legislature to consider its legal and ethical implications and effects, without stifling innovation. However, this technology can generate tensions concerning fundamental rights, legal security, or with the principles and obligations of the public administration (Cotino Hueso 2017; Thelwall 2018; Valero Torrijos 2020).

The interest of states and supranational organisations in this matter has been reflected in declarations, guidelines, plans, and strategic programs to integrate AI in contemporary societies to encourage its progress and, at the same time, safeguarding possible interests and rights. In the Spanish case, the development of intelligent administration has been uneven, although in certain areas the use of AI is advancing faster, such as the case of police investigations or in tax and social security matters. Current Spanish regulations have established electronic administration and support the expansion and implementation of new technologies, but they may not be adapted to the legal needs generated by AI. This occurs despite having several projects based on AI, such as the one developed by "Consejería de Empleo, Formación y Trabajo Autónomo de la Junta de Andalucía" to reduce the time of resolution of subsidies in the context of the COVID-19 pandemic for the self-employed. Furthermore, similar tools such as "data mining" are being used by public administrations to analyse large amounts of information to extract patterns or identify relationships within the analysed data set to formulate predictions. These are used to determine the probability that irregularities have been committed in a public procurement procedure (Criado 2016; Cerrillo i Martínez 2017).

The goal of this article is therefore to define what are the risks derived from this new technology that Spanish public administrations may face and if the current law has mechanisms to deal with them. For this reason, the research questions will be:

- Do we need a legal definition of AI?
- What are the risks associated with the AI uses in Spanish public administration?
- Can current Spanish legal mechanisms solve them?

The initial assumptions are, on the one side, that the definition of AI is vital for regulation and governance because laws and policies need it to operate and, on the other side, that the use of AI can generate problems such as gender biases or intrusions to the privacy of citizens and that Spanish legislation is not prepared to provide a solution. Based on the existing intelligence administration and AI literature, a semi-structured interview guide for the expert interviews has been derived. Fourteen interviews were conducted with experts knowledgeable about the development of AI and its applications in public administration. The experts included researchers from prestigious Spanish and foreign universities, IT public workers, public managers on the national, regional, and municipal government levels, and experts from the legal profession. The interview guideline addressed topics such as the existence of gender bias in current AI projects, the lack of transparency in the use of this technology, or the legislative possibilities to fight these risks.

In order to answer the research questions, a dual structure was followed. On the one hand, the first part is analytical and descriptive, which allows the setting of the foundations of the concept of AI in public administrations, also its use and legal guarantees both in the European Union and in the Spanish case. This theoretical framework will be completed by an overview of the benefits and risks of the implementation of artificial intelligence in public management. On the other hand, the second part of the research focuses on the analysis and discussion of the interviews carried out.

In the following, the theoretical framework is provided for the appearance and regulation of artificial intelligence in the public sector at European and Spanish levels and its risks are presented. Then, the methodology used and analysis steps are described. Finally, in an inductive process, the findings are derived and are discussed in the context of the existing literature. The paper concludes with a conclusion and a set of limitations of this research.

2. Background

AI is having a great impact in contemporary society; its use in the private sector is widely extended. For example, algorithms are being used to improve computational language (Nowicki et al. 2021) or for the development of autonomous cars (Harris 2018). However, the value creation and functioning of AI in specific public uses are also evident. Lately, Wirtz et al. (2019) identified a set of types of AI uses in public services, such as predictive analysis, simulation, and visualisation of data to prioritise those areas that require inspection by administrations. Consequently, Todolí Signes (2020) declares that automated tools can reduce the time and resources used, and also determine non-compliance patterns and trends that would be undetectable by human experience and intuition. Moreover, these instruments help efficient planning in the medium and long term. In this sense, the benefits that new technologies can bring to public administrations are many, including the possibilities in public procurement with the use of techniques for cross-analysis of huge amounts of data (Valcárcel Fernández 2019; Sobrino-García 2021a).

2.1. Theoretical Frameworks on Artificial Intelligence in the Public Sector. Concept, Benefits, and Risks

The notion of AI is related to computer systems capable of thinking, learning, collecting data and information from multiple sources, and acting according to several objectives correlated to algorithms. The creation of these algorithms is a statistical, mathematical, and also human process, including the large amount of data collection and analysis in different phases (Coglianese and Lehr 2017). According to the European Commission (2018), AI refers to systems that display intelligent behaviour by analysing their environment and taking actions, with some degree of autonomy, to achieve specific goals. Therefore, AI-based systems can be software-based, in the case of voice assistants, or AI can be embedded in hardware devices, such as autonomous cars.

Deep learning and machine learning are the subfields or applications of AI, which deal with designing algorithms capable of educating machines by helping them recognise patterns and extract knowledge from previous cases (Ali and Frimpong 2020). An example of the types of algorithms that use deep learning are those used to calculate the risk of recidivism or in the creation of targeted advertising adapted to a specific audience (Murphy 2012; Tegmark 2017). Deep learning is inspired by the functioning of neural networks in the human brain, and the data goes through different “layers” in which learning rules are applied (Bertolini et al. 2021). AI techniques begin from a set of input and output variables. The relationship between them is established through a training or learning process carried out by algorithms, guided by large amounts of data. This system requires a sequence of instructions that specifies the actions to be executed by the computer system (Navas Navarro 2017). These algorithms can build solutions or AI models for certain problems. Usually, an AI algorithm assigns the same relevance to input variables such as a person’s income level, zip code, or ethnic origin, unless the creator indicates otherwise. If these variables allow the performing of the objective of the algorithm, they will be considered

important. Therefore, an algorithm does not seek to certify the veracity of a hypothesis, but rather to look for correlations between the data (O'Neil 2016).

This is precisely one of the keys to algorithms: their advantage lies in their ability to anticipate behaviours, guess trends, or witness plausibility. Algorithms could support operational management and service delivery by public organisations. For instance, they could enable organisations to deploy people and resources in a highly targeted way when undertaking audits. Moreover, the technology underlying an algorithm enables decision-making processes to be made more transparent and easier to audit (Valle-Cruz et al. 2019; Zheng et al. 2020). Nevertheless, the constructions generate correlations not connections of causality and, in general, they are unable to provide the codes to their operation, with the only exceptions of some particularly sophisticated algorithms. Consequently, this system presents several problems: the models can generate biased predictions (Caliskan et al. 2017) or that the use of numerous variables combined with complex algorithms could lead to models incomprehensible for humans (Zlotnik 2019).

Although the principal motivation for the use of this technology was to solve problems objectively, in many cases the algorithms or the data set used by them may contain certain biases that lead to discrimination. People also have some in-built biases, but there is a risk in using an algorithm that it may be primarily dependent on decisions taken by the programmer or data scientist (Noble 2018). The discrimination resulting from algorithmic processes is produced by the introduction of data processing and decision automation technologies, although they also present problems of opacity, difficulty in assigning responsibility for decisions, and risks to the intimacy and privacy of people (Soriano Arnánz 2021). Finally, some algorithms used by public administrations or central governments have been obtained from external companies. The data and mechanism used by these algorithms are often owned by the external company, who may wish to protect this information. Where liability or aspects such as the processing of personal data are concerned, the public administration cannot simply rely on the information provided by the company. This makes analysing and managing the risks associated with the algorithm more difficult for the public sector. For this reason, the use of AI by public administrations must be guided by the guidelines of "good administration" (Ponce Solé 2019), understood as the principle according to which public administrations correctly perform their purpose, serving the general interest, and making an adequate weighting of the means, circumstances, facts, and elements present (Menéndez Sebastián 2021). In other words, the use of AI by public administrations must serve the general interest of citizens, with respect for private interests and guaranteeing fundamental rights. Therefore, the public sector literature has considered the adoption of soft law criteria as beneficial for the development and implementation of AI in public administrations (Sarmiento 2008; Valero Torrijos 2020), that is, the adoption of agreements, policies, or codes of conduct that establish performance standards.

2.2. Policies on AI in the European Union. Ethics and Human Rights as the Core to the Next Regulation

According to the previous literature, AI allows the management of large databases to improve the work of public sector professionals, especially from the integration of internal and external databases, even if they include information of a different nature, and both quantitative and qualitative, to generate new results. For this reason, different initiatives have emerged in the European Union (EU) that promote AI development strategies in the public sector, as this technology has proven to be a powerful transformer on the European economy and competitiveness. Precisely, the "Digital Agenda for Europe", are a set of policies that pretend to redesign the European telecoms sector and to achieve a "data-driven" public administration based on technology such as AI, blockchain, or big data. Still, the EU has also been concerned about the risks involved in this technology and has overseen the promotion of a framework based on European principles and values. Consequently, the EU is attempting to create a policy on AI and an incipient structural legal framework to support it. The European Parliament (2019a) highlighted not only the benefits of AI in public administration, but also a number of ethical, legal, and economic

concerns relating to the risks facing human rights and fundamental freedoms. Thereby, a human-centred AI should be human rights-based (Raso et al. 2018). Even though there has been some limited discussion at the European government level of the impact of AI on human rights, especially regarding the right to privacy, the impact on social and cultural rights has so far received little attention (Fernandez-Aller et al. 2021).

However, some authors declare that jurisdictions that understand the limitations of machine inferences that feed on machine readable human behaviours will gain a competitive advantage compared to jurisdictions that fail to consider the human-centred AI (Hildebrandt 2020).

Nowadays, several ethical codes and guidelines for AI development have been drawn up by different organisations, frequently in response to growing awareness of the possible adverse effects of this technology (Morley et al. 2020). Until now, policy documents related to ethics and AI have also focused more on ethical frameworks than on possibilities for enhanced regulation (Vesnic-Alujevic et al. 2020). Since 2017, the EU has been trying to establish a common policy on AI, at the same time the Commission was mandated to analyse the risks of AI to guarantee privacy and ethical standards. Next, in the communication “Artificial intelligence for Europe”, the fundamental elements about this technology were established. These can be compiled into three ideas: first, the adoption of measures in the educational and labour spheres to face the economic changes that came from the fourth industrial revolution; second, the promotion of technological and industrial capacity using AI; and finally, the development of an ethical–legal framework based on EU values. These standards should be elaborated according to other European standards, such as data protection, cybersecurity, or open data.

Afterwards, the European Commission and the member states signed a declaration on cooperation in the field of AI, which includes the intention to cooperate in this area and to establish policies at a national level. Furthermore, an action plan was elaborated, which promoted coordinated and harmonised actions linked to the development of an ethical framework and the approval of technical standards, for lack of European regulation. At the beginning of 2020, the Commission published the “White paper: on Artificial Intelligence—A European approach to excellence and trust”, which presents several policy options to enable trustworthy and secure development of AI in Europe, in full respect of the values and rights of EU citizens.

The objective of the white paper was to promote different regulatory proposals, highlighting the urgency to develop an AI ecosystem focused on citizens, business development, and services of public interest. Later, the European Parliament approved three reports that analyse how to regulate AI according to respect for ethical standards and trust in AI technology. These reports had three basics: first, the balance between the protection of rights and the technological impulse; second, the existence of a civil liability system; and finally, an intellectual property system, as well as guarantees for developers. These three elements must be followed by a regulation that supports human intervention and supervision of this technology. In short, although the EU is beginning to generate the bases of an AI regime, it does not have a coherent legal framework that helps to harmonise the use of this technology.

Although, since 2018 with the Coordinated Plan on Artificial Intelligence, several European countries have developed their national AI strategies. Currently, more than 20 countries have published national AI policies (Van Roy et al. 2021).

The European Proposal for a Regulation on the Uses of AI

After all the communications, reports, and the white paper, the Commission published in April 2021 a proposal for a regulation to harmonise rules on AI. This project defines the AI system as a software developed with one or more of the techniques and approaches listed in an annex and can, for a given set of human-defined objectives, generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with. This annex contains a detailed list of approaches and techniques for

the development of AI to be adapted by the Commission in line with new technological development. Specifically, it has set three different approaches: the machine learning, the logic- and knowledge-based, and the statistical methods. The main basis of this future regulation is to take advantage of the benefits that AI offers and, at the same time, prevent the risks that can derive from its use. These risks involve the possibility of people's integrity or life damage, serious impacts for society or economic activities, or even negative impacts on fundamental rights (Fernández Hernández 2021).

This regulation proposal combines different intervention techniques from the total or partial prohibition of certain activities to avoid risks, through preventive controls with authorised regimes, to an inspection and liability system. Apparently, this proposal prohibits some uses of AI, specifically those that manipulate citizens, those that use personal information to detect vulnerabilities of personal ones, and those that use personal information to establish classificatory profiles of citizens. Although some of these uses may be performed by a public authority in exceptional cases authorised by a rule and for public safety purposes, it seems rather that this exception is intended for generic surveillance (Huelgo Lora 2021). On the other hand, there are a group of applications that require authorisation, such as remote biometric identification in public places, which will require administrative authorisation if there is an enabling regulation. Moreover, specific rules will be established for certain cases; for example, when the public administration or the entity use automated mechanisms or chatbots that interact with users, they will be warned that they are not talking to a real person.

However, the true core of the European proposal is the approach to the uses of AI based on the risk analysis that this system may generate for the fundamental rights and security of people. The proposal includes a series of high-risk applications that will require specific conditions; for example, in the case of biometric identification, a prior verification by an independent third party is required, as well as, for the operation of critical infrastructure, that is, systems whose malfunction can cause very serious damage, such as a power station. This variety of AI systems will be subject to severe obligations, having to guarantee that they have risk assessment and mitigation methods, high quality of the data used in the system, comprehensive documentation on its development and purpose, and human supervision to minimise the risk. Nevertheless, the European legislator has given some flexibility to the requirement of transparency to respect the industrial secrets used in the application. At the same time, those applications that are not classified as "high-risk" may assume voluntary codes of conduct. Moreover, member states will have to establish supervision and sanction mechanisms. One of the last issues that the proposal regulates is the establishment of sandboxes or controlled test spaces.

In short, this regulation proposal establishes some requirements which allow for AI applications to be used by both public and private operators. They are general and additional requirements that seek to prevent damage to protected property and rights at the highest legal level. In other words, the proposal tries to prevent those harmful results, not forbid things that are now allowed.

2.3. AI in Spanish Public Administration. The Legal Guarantees of Administrative Law

The European commitment to the implementation and development of AI in public administration has received significant support from member states. In the Spanish case, the approval by the Government of the "Spain 2025 Digital Strategy" and the "Spanish Strategy in R&D in artificial intelligence" are aligned with the Digital Agenda for Europe to achieve a public administration at the forefront of new technologies. For years, the framework of the Spanish digital administration model has as its principal axis the adoption of new information and communication technologies (ICT), intending to promote the so-called information society. Thus, national strategies for digital transformation were produced during the first decade of the 21st century and focused on the development of digital public services. Since then, a series of principles focused on transparency, citizen participation, and public innovation have been addressed (Criado 2016), and consequently,

the use of ICT was expanded within Spanish governments and public administration. Along with the political-strategic dimension of the Spanish digital administration is the regulatory and normative component. At the beginning, Law 30/1992 on the Legal Regime of Public Administrations and Common Administrative Procedure contemplated the possibility of a future electronic provision of public services, defining common bases and general rules for the risks that could arise in the use of new technologies in public administration. According to successive modifications, key aspects of security and privacy were included, as well as an ICT management model in public administration. Next, in 2010, two royal decrees came into force, the National Interoperability Scheme and the National Security Scheme, which, joined to the associated technical standards, established the basis for the exchange of data between public administrations under the necessary security standards. However, it was not until 2015 that Law 39/2015 of the Common Administrative Procedure of Public Administrations and Law 40/2015 of the Legal Regime of the Public Sector came into force. These regulations attempt to guarantee the right and obligation of citizens to interact with public administrations through digital mediums, from the simplest administrative procedures to identification systems and electronic signatures. Particularly, Law 40/2015 established a regulation on the use of electronic means in the internal functioning of the public sector, opening the door to automated administrative action. Consequently, for the first time, the possibility of using technologies such as AI in the Spanish public sector was included. Nowadays, this legal system includes the possibility that public administrations can adopt automated decision-making using the electronic means available. Nevertheless, the use of AI in public decision-making requires the consideration of certain elements such as the own connection of the decision with the public administration, the phases of the administrative procedure in which the algorithms may intervene, or the motivation of these public decisions (Cerrillo i Martínez 2020a). Still, it is a limited and dispersed regulation, with gaps in the procedure for the creation of the algorithm and the explanation of the automated decisions. Despite these limitations, this is the only administrative-legal provision related to AI's particularities (Ponce Solé 2018).

The governments of the Spanish autonomous communities have already begun to implement AI projects in the public sector. For example, the Comunidad Valenciana has incorporated a security administration tool for network analysis (known as SATAN). This software cross-databases the public administrations to identify corruption cases in administrative processes. SATAN analyses millions of data in real-time and allows for alerts about possible conflicts of interest; for instance, if there are irregular extensions or modifications of the public contract's specifications after awarding them, or even if several offers have been obtained in the processing of a contract. In these cases, when the system detects signs of corruption risks, it categorises the alarms according to how important they may be and sends an alert to the inspectors in charge of analysing the information.

In short, in Spanish administrative regulation, there is neither definition of algorithms, nor a legal regime that guides the application of algorithms in the public sector. However, the absence of a specific regulation does not exempt the application to the rest of the administrative law regulations, as well as the principles of transparency, open government, and good governance (de la Sierra Morón 2021).

3. Materials and Methods

The research questions were designed to investigate the phenomenon of AI in public administration in the Spanish case, to determine if an AI definition is required, and to understand the risks that may exist in the public sector and verify if the Spanish legislation can mitigate these threats. Therefore, a qualitative and interpretative stance was adopted, aiming to understand the approaches and perceptions in the uses of AI in Spanish public administration.

As the method of inquiry, we chose semi-structured interviews with experts to collect data directly from those subjects involved in AI projects with a broad overview of public administration decisions and in-depth insights about implementation actions. The semi-

structured interview is a qualitative data collection strategy in which the researcher asks informants a series of predetermined but open-ended questions. This method consists of a dialogue between researcher and participant, guided by a flexible interview protocol and supplemented by follow-up questions, probes, and comments. The method allows the researcher to collect open-ended data, to explore participant thoughts, feelings, and beliefs about a particular topic and to delve deeply into personal and sometimes sensitive issues (Dejonckheere and Vaughn 2019). An expert is defined as a person with technical, process, and interpretative knowledge in relation to their areas of expertise. They have more than just systematic organised knowledge—experts also have deep knowledge in specific experiences which result from their actions or obligations of the specific functional status within an organisation (Bogner et al. 2009). We opted for expert interviews because of the viewpoints they provide and for their privileged access to public policy-making processes. Specifically, for the semi-structured interview questions, we followed different thematic guidelines, but they did not have a specific wording or exact order (Martínez 2006; Valles 2009).

In this research, an interview analysis approach was used that focused on meaning, according to four stages (Silverman 2013; López-Chao et al. 2019). First, the content analysis, which implies the codification of the meaning to facilitate its categorisation, and for this reason, the analysis codes were established. Second, the encoding of meaning, which implies giving a keyword to a specific part of the text to facilitate its identification. Third, the meaning condensation; in this stage, the meanings expressed in short formulations were summarised. Finally, the interpretation of the meaning; in this phase, we proceeded to the interpretation of the text. This type of analysis allows for attention to the linguistic features of an interview, being able to generate and verify the meaning of the statements, as well as to improve the precision of the questions in the interview (Qu and Dumay 2011).

3.1. Data Collection, Procedures, and Instrument

To ensure precision and rigour, the sample selected was defined by its ability to represent salient characteristics and features of relevance to the investigation. Consequently, the first step was to decide about the sample by those who were able to provide the most relevant information. Therefore, we selected different personalities from the legal, academic, and technical fields who worked in the elaboration, implementation, and regulation of these types of algorithms in Spain, intending to know practical experiences. Experts were involved directly in AI and digital transformation projects in public administration and presumably exposed in their real-life settings to the central phenomenon of this research. This method allowed us to generate the data necessary to determine the different categories of the main concept from the experts' perspective and to delineate any differences (Ragin 2009).

The selection of the experts was based on their national status in the area of AI in public administration. The experts selected are publicly known to have high-level overview of the topic, are known to have made statements and research projects about the legal guarantees of AI in administrative activity, and have special knowledge and experiences based on their functions or responsibility. The experts came from different public administrations (e.g., universities, councils) or other organisations involved in AI public sector projects (e.g., IT service providers, consultancies).

In total, 14 experts from Spain were interviewed. Eight experts were academic professionals who came from public universities and four experts who came from private organisations, including consultancies that specifically advise governments and several public administrations. Moreover, these organisations are specialised in supporting or providing services in the public sector. The last two experts came from government (ranging from the national level to municipal level). Academic experts have been part of research projects (at the national and municipal level) in which possible difficulties related to the use of AI in different areas of the public sector have been studied. The private sector experts are from legal consultancies which advise public administrations on AI policy and

its implementation. Although these organisations are from the private sector, the experts were able to provide relevant perspectives to this research.

The expertise of the 14 interviewees included implementation, strategy, or policy development related to artificial intelligence. Some experts were involved in public service automation projects linked to machine learning. Other experts described their participation in the development of policies for the national and municipal level. The interviewees from consultancies gave legal assistance in the development of policies or the implementation of public sector policies.

The prosperity of the research based on semi-structured interviews with experts depends on the number of interviews conducted, the quality of the experts interviewed, and the quality of the experts (King et al. 2019). According to the previous literature (Glaser and Strauss 2000), at least ten interviews are required to adequately analyse the patterns and differences across subjects. Therefore, fourteen interviews were conducted with experts knowledgeable about the development of AI and its applications in public administration. The interviews were conducted using online meeting tools and were transcribed into word files, eliminating the personal data of the interviewees, and all the interviews were conducted in Spanish. Thus, some of the nuances of the language may have been lost during the translation for the transcripts. This modality was chosen due to the problems caused by the COVID-19 pandemic. For this research project, the interviews were conducted between September and October 2020 by contacting the experts using an email template to contact and invite them.

Semi-structured interviews included a short list of guiding questions that were supplemented by follow-up and probing questions that are dependent on the interviewee's responses. All questions should be open ended, clear, and avoid leading language. Consequently, for this research, we developed some thematic guidelines linked to the research questions to guide the interviews. Following these guidelines, we formulated several standard questions to introduce the subject (Table 1).

Table 1. Semi-structured interview and standard questions.

Semi-Structured Interview and Research Questions		
Research Questions	Thematic Guidelines	Standard Questions
Do we need a legal definition of AI?	Concept of AI and the trust of citizens	How do you define AI? Can we trust AI? How can AI improve the public sector?
What are the risks associated with the AI uses in Spanish public administration?	Possible risks in the use of AI	What are the main risks for the public sector? What effects does AI have on privacy? Can AI technology lead to biases?
Can current Spanish legal mechanisms solve them?	Compatibility with current legislation	How should AI procedures be regulated? How to update the legislation of artificial intelligence so that it is fairer and considers its risks? Do we have legal and administrative mechanisms to regulate AI?

3.2. Data Analysis

For the data analysis, we used a common approach based on using codes from a codebook for tagging segments of text and then sorting text segments with similar content into separate categories for a final distillation into major themes (Table 2). This approach has been described as a 'template approach' as it involves applying categories based on prior research and theoretical perspectives (Miles 1994). Specifically, there are three elements used in the analysis of the information in attention to the analysis processes: data reduction, provision and transformation of information, and extraction and verification of conclusions

(Ballester et al. 2003). These elements have been used to facilitate the identification of the fundamental ideas of the interviewees, considering the different factors that affect the subject. Later, we used an analytical process to reconstruct them in a structured and meaningful way, to obtain a series of conclusions to carry out the subsequent triangulation of results.

Table 2. Analysis of the interviews for stages.

Analysis of the Interviews	
Phases	Explanation
Phase 1	During the first cycle of the data analysis the coding and categorisation of the information was applied.
Phase 2	To divide the information into categories, an analysis based on thematic criteria was followed, in which the text was reduced according to the criteria addressed.
Phase 3	In the third phase, we made a grouping that allowed us to generate a structure of the information. Next, we transferred this information into a text document to facilitate its visualisation and presentation.
Phase 4	Finally, we elaborated an assessment that allowed us to conclude the transmitted experiences, existing patterns, or generalisations in the investigated field.

The analysis of the interview resulted in four categories (Table 3). Three of them directly related to the research questions: the concept of AI, risks of the use of AI in the public sector, and legislative and administrative instruments. The fourth category, “transparency and trust in AI algorithms”, emerged as a concern among those interviewed due to the importance of the transparency obligation in the Spanish public sector.

Table 3. Summary of the categories and indicators of AI in public administration.

The Use of AI Technology in Spanish Public Administration	
Categories	Indicators
Concept and design of the AI algorithms	autonomy, intelligence, human process, deep learning, machine learning, training
Transparency and trust in AI algorithms	open or semi-open code, neutrality, human intervention
Risks in the use of AI algorithms	biases (gender), privacy, opacity, algorithms’ legal status
Legislative and administrative instruments	ethical principles, control mechanisms, audits, protocols, collegiate bodies, soft law

4. Results

The results in this section are presented from the categorisation and analysis of the transcripts of the 14 interviews. We organised our findings along the four dimensions derived from the coding and related to the research questions. Results of the data analysis approach are displayed, mainly using quotes from the interviews with the experts. We present the results obtained according to the categories to establish the bases that, together with the theoretical framework, will answer the research questions in the following section.

4.1. Concept and Design of the AI Algorithms

The previous literature has established a concept about AI according to several characteristics (e.g., autonomy). Furthermore, different European organisations have contributed through their guides and documents. However, the interviewees agree that a harmonised definition should be established as a starting point.

“We need a homogeneous and clear definition of artificial intelligence. Computerised systems capable of thinking and learning by themselves. The creation of these algorithms must be understood as a human process.” (Interviewee 10)

This technology has numerous definitions given by consolidated research. However, and as the interviewees point out, we can take some characteristic features: computer systems, learning capacity, algorithm training, and autonomy. AI tries to emulate the capacities of the human brain, and for this reason, this technology is potentially applicable to any field of human intellectual activity.

“An AI system needs a sequence of instructions that specify the different actions that the computer must execute to solve a specific problem. This sequence of instructions is the algorithmic structure used by the AI system. Therefore, “algorithm” is the procedure to find the solution to the problem.” (Interviewee 4)

Parallel to that, the interviews have highlighted that AI is only one of the automation systems since there are other methods with algorithmic techniques. There are machine learning and deep learning, allowing the development of an approach with a data orientation to educate machinery to promote its autonomy.

“A system with machine learning could learn the data and classify with greater precision, but one with deep learning can ‘train’ with the new data it receives. That is, it can use the wrong differentiator and make a mistake once, but the next time the system use another it gets closer and closer to the correct result.” (Interviewee 8)

Next, another theme that came out in the semi-structured interviews was the design of algorithms for AI. According to the interviewees, in the design of these systems, the ethical approach must prevail and be implemented in the business, professional, and technological sectors, and even in promoters of public policies and future regulations. The possibilities that AI offers raises a series of questions that go from the own ethics of researchers, designers, and users to the moral behaviour of the artificial products resulting from the design. The insertion of an ethic in the development process can promote or force the organisations to structure and adapt to regulatory compliance, generate governance that is sensitive to problems, and insert various internal and external control systems.

“Certainly, an AI design based on ethics can go a long way toward introducing positive values into society and preventing or correcting injustices.” (Interviewee 14)

In this sense, several interviewees highlighted that the proposals of the European Parliament on the ethical framework for the design and production of this type of technology are an essential step since they imply respect for human dignity, autonomy, and self-determination of the person.

“We also need to move forward in ethics, perhaps with an ethical statement for AI professionals like the one used by other professions (lawyers, doctors).” (Interviewee 1)

4.2. Transparency and Trust in AI algorithms

Apart from the elements related to the research questions, another concern that arose among the interviewees was trust in this system and operation transparency. In general, experts agree that we should trust AI since it can help find solutions to many of society's problems, as we remarked in the background of this research.

“We can trust in artificial intelligence, but we need to improve some issues related to the transparency and comprehension of algorithms and human supervision of all processes.” (Interviewee 1)

However, some experts agree that the problem lies in the human capital, which works behind the AI design. They expose a fascinating and controversial idea, and that is trust in the human being. Cultural or ideological values pervade people, which can lead to biases in algorithms. Moreover, two experts point out that one of the problems regarding AI mistrust is the understanding of the deficiency of the algorithm's operation, especially between public administrations.

“The question would be if we can trust humans. Artificial intelligence is a development of the “human” for this reason, it carries several “pros and cons”. Yes, we can trust

artificial intelligence but—like everything else—the level of trust will depend on where it comes from. For example, is a system developed in the European Union the same as one in China? No. Since they do not have an equal legal system. This situation is the same for companies, some are more reliable than others.” (Interviewee 4)

In the face of the possible problems of lack of trust, elements such as transparency appear, in relation to dimensions such as how an algorithmic decision is arrived at and based on what assumptions, and how this could be corrected to incorporate feedback from the involved parties. The process of algorithmic auditing within the software development company or public entity could help in tackling some of the ethical issues raised. Technology cannot advance outside of society, and there is the base of trust: appropriate and effective control.

“Elements such as transparency (what data it uses, what decisions it makes, etc.) are essential to monitoring the performance of the algorithm once it is in a state of implementation. The possibility of carrying out external audits, for example, would also be interesting to study.” (Interviewee 6)

4.3. Risks in the Use of AI Algorithms

The third category includes the conflicts that we identify both in the theoretical framework and in the interviews with experts. Specifically, we divide them into four subcategories: algorithms' legal status, gender biases, privacy, and opacity.

Algorithms' legal status. Regarding the first of the risks, some interviewees identified that one of the conflicts derived from algorithm uses in the public sector is their legal status. One of the interviewees points out that the Spanish public administrations have denied access to the algorithm in certain cases, due to the uncertainty about its status as information accessible to the public. Nevertheless, this case has been resolved with the statement that the algorithm is also public information if the algorithm is held by the administration. Furthermore, the interviewee also remembers the decision of the *Tribunale Amministrativo Regionale* of Italy, in which it is considered that an algorithm used by the Ministry of Education was an administrative act.

Biases/Gender biases. Another of the risks detected was the existence of biases. In this sense, the interviewees highlight that machine learning can cause disproportionate damage to minority and oppressed groups, showing some concern about gender biases.

“The biases depend on the data that the algorithm uses, that is, if the data includes this type of discrimination, the biases are transmitted to the algorithm. These can be produced by the prejudices of the designers.” (Interviewee 5)

A major driver of bias in AI is the training data. Different types of massive data can be used for algorithm training; on the one hand, those that arise from the interaction of people on the Internet, for example, in social networks or internet forums, and on the other hand, the data generated by machines and technological devices, such as those obtained by GPS sensors, electricity, or water consumption data, etc. Due to this fact, there is increasing concern about the existence of deficits in the data quality, particularly in the gender biases related to the underrepresentation of women and specific groups in the databases. However, some interviewees opined that the biggest problem is the perpetuation of certain roles and biases in the social consciousness.

“Whoever designs the algorithm can determine the existence of biases in the algorithms. This can be conscious or unconscious, reproducing certain social patterns that may involve biases. However, although the design of an algorithm by a woman can avoid some of these biases, it may not happen in all cases since certain patterns are unconsciously maintained.” (Interviewee 2)

Nevertheless, Interviewee 13 reflected the need to treat all biases with the same relevance, bringing up the algorithm's cases used for image or voice recognition based on gender stereotypes according to people's physical appearance. These algorithms reproduce

false assumptions about how people's physical attributes should be according to their biological sex and ethnic or cultural origin. Public administrations are guarantors of equal rights and opportunities for citizens, and because of this, the existence of any of these biases in the algorithms that public administrations may use is unacceptable.

"I believe that biases, all of them, have to be dealt with holistically. Besides, the problem of bias must be approached from the design, and in AI implies the analysis of the data used for development, and of the variables selected in these in the first place." (Interviewee 13)

Privacy. A third risk that generates a big concern among citizens is related to artificial intelligence's impact on personal data protection and people's privacy. These doubts derive from the fact that the current Spanish administration has access to and generates a large number of citizens' data. For this reason, it must be considered that the processing of personal data in public administrations involves different risks compared to others and that they derive, at least, from the volume of affected subjects, from the extension of the data collected, or from the impossibility, in many cases, to the objection to processing personal data. Although the public administrations guide their data processing by a spirit of public service, various risks could materialise for citizens in certain situations. For instance, these include situations of the government of law bankruptcy, the public administrator's abuses, in circumstances of massive personal data or selective filtering because of security breaches, or in the event of legislative changes that affect cross-data transfers. Therefore, the interviewees remember that the development and design of AI projects must follow current data protection legislation.

"When the algorithms process personal data, the adoption of the necessary measures for the management of privacy and personal data is required under the provisions of The General Data Protection Regulation (GDPR)." (Interviewee 2)

Opacity. Good administration requires transparency, although when public administrations use AI there is a risk of opacity. In fact, it is common to refer to machine learning algorithms as black boxes.

"The opacity or lack of transparency is one of the main characteristics of the algorithms (black boxes) and one of its principal defects. To face this undeniable reality, major technology companies have also begun to recognise the problems that come with this type of AI technology." (Interviewee 12)

The interviewees commented that the algorithms could be opaque for different reasons: for technical reasons (machine learning algorithms are dynamic and complex, which makes it difficult to know their design), for legal reasons (the legislation on transparency and access to public information recognises several limits to avoid harming certain legal assets), or for organisational reasons (it is common for public administrations not to formalise decisions regarding the use of algorithms).

"Algorithmic opacity makes it extremely difficult to detect algorithmic bias. Principles such as the explainability of the algorithm, its fearless face enormous difficulties in achieving their objectives due to the extremely high complexity of many algorithms, even for technological experts." (Interviewee 3)

Along with the complications perceived by Interviewee 3, for his part, Interviewee 2 remarks that the algorithm used by public administrations is public information, and therefore, the transparency mechanisms included in Spanish legislation must be followed.

4.4. Legislative and Administrative Instruments

The interviewees gave a unanimous opinion that the current Spanish legislation is not prepared to provide a solution for AI problems. For this reason, it is necessary to develop administrative and legal mechanisms. Therefore, they propose various alternatives ranging from the creation of collegiate bodies or committees to control the progress of AI projects to the development of ISO certifications. The latter refers to a certificate that verifies that an

entity complies with a set of regulated guidelines and is in accordance with the reliability standards in the regulated matter.

“Control over AI programming is essential and must be carried out by independent bodies or mechanisms, preferably collegiate and with the participation of the people/groups that may be affected by the decision. However, control/supervision cannot be limited to the initial moment, prior to its operation, but, on the contrary, it must be projected in an evolutionary way to the later phases of system operation, particularly if non-deterministic algorithms are used.” (Interviewee 8)

Other interviewees opined that the key lies in the adoption of principles with the human-centric approach, the values of respect for human rights, dignity, freedom, democracy, and equality. The EU and Spain have a solid regulatory framework that will set the global standard for human-centric AI. In addition, the General Data Protection Regulation ensures a high standard of personal data protection and requires the implementation of measures to ensure data protection by design and by default. For this reason, Interviewees 6 and 9 propose the implementation of a prior impact assessment, such as the evaluation contained in the GDPR, to verify the development of the AI algorithm according to the standards and principles set.

“We can take advantage of AI by adapting it to principles such as loyalty, explainability, or accountability. Different principles focused on achieving AI person-centred and therefore, with an egalitarian approach. In this sense, technological, ethical, and training measures can be implemented, especially the latter aimed at software designers.” (Interviewee 12)

“The quality of the artificial system could be assessed with a prior impact assessment (article 35 GDPR) but including equality. This evaluation should be carried out into the administrative procedure that leads to the approval of the algorithms if it is designed by the public administration itself. Although, in those cases in which the development of the algorithm is contracted to a private company, this evaluation must be agreed in the contract.” (Interviewee 9)

5. Discussion

This study has attempted to answer three research questions: (1) Do we need a legal definition of AI? (2) What are the risks associated with the AI uses in Spanish public administration? (3) Can current Spanish legal mechanisms solve them? In this section, we determine the need for a legal definition of AI and give the defining characteristics. Here, we provide an empirically based description of the principal risks of AI uses in Spanish public administrations and we contribute with possible regulatory proposals to minimise or prevent these difficulties. Furthermore, the results from expert interviews evidenced another element, the trust and transparency in this technology use.

In this section, we discuss how our findings extend the existing literature on the AI technology uses in public administration and its risks. Then, we also discuss how our findings provide interesting regulatory proposals not only for the Spanish case but also to be extrapolated to other countries with a similar administrative system.

5.1. Concept, Transparency, and Trust of Artificial Intelligence Algorithms

The definition of AI is diverse in the existing literature, and therefore, the characteristics emphasised will fluctuate a little depending on whether the research comes from fields such as computer engineering (Abdeldayem and Aldulaimi 2020; Batarseh et al. 2021) or others associated with the social and legal sciences (Raso et al. 2018; Campos Acuña 2019). However, the experts interviewed have pointed out common characteristics of AI, such as autonomy, learning capacity, algorithm training, and the fact that it is a computer system. These features have been supported by the existing literature (Cotino Hueso 2019; Ali and Frimpong 2020). In response to the first research question, having a precise legal definition accepted by the scientific community will avoid ambiguities and legal uncertainties. For this reason, the above characteristics are crucial to developing

a solid and homogeneous legal concept. Actually, the need for a legal definition of AI comes from the need to legitimise it, that is, to establish it according to the rules of law and with the aim to unify the concepts to avoid legal loopholes. The European Union has tried to establish a definition through its working documents, a starting point for the adaptation of the countries that compose it, including Spain. This term contains an explicit reference of the notion of intelligence, nevertheless, since intelligence is a vague concept the notion of rationality is better. This refers to the ability to choose the best action to take in order to achieve a particular goal, given concrete criteria to be optimised and the available resources. Our results have made it possible to extract the principal characteristics of AI, supported by the existing literature, to establish a base for the Spanish legislator to adopt an AI definition.

The approval of a legal term for this technology will also transmit trust among citizens, another of the concerns underlined by the experts interviewed. Thus, the design of the artificial intelligence algorithms used by the public administration has been another question debated in the results, prevailing the need for following an ethical approach in both the private and public sectors. Most of the experts affirmed that the characteristics that an algorithm must have to guarantee transparency and provide trust were that it had an open-source code and training with open data. This type of algorithm would be auditable and, therefore, could be subject to citizen control (Vesnic-Alujevic et al. 2020). However, some authors differ on the previous affirmation. The reason behind this is related to the importance of the intellectual property rights of the company algorithm (Ponce Solé 2019). The protection of the source code seeks to avoid its publicity, and for this reason, its approach as open source would be complicated if the development of the algorithm is made by a third party and not the public administration.

From the results of the interviews, we were able to extract another proposal for the adoption of an ethical approach in the algorithm's elaboration. That is, the implementation of an ethical statement for AI professionals like the existing statements in other professions, such as the Spanish Legal Professional Code of Ethics. This code would ratify and elaborate behaviour standards to respect the higher values of society and human rights in the development of algorithms. The norms dictated in the code of ethics would be previously agreed upon and approved commonly and unanimously by all the members of the profession. These would be guidelines of conduct to carry out an adequate job and help the whole of society that requests the services of the profession to obtain full satisfaction with the proper execution of the work.

5.2. Main Risks of Using AI Algorithms in Public Administration

In response to the second research question, we identified four principal risks in the algorithm uses by Spanish public administrations: the biases (gender), the privacy, the algorithm's opacity, and the algorithm's legal status. One of the main concerns detected in the results was the possible existence of discriminatory biases, with significant relevance of gender. This is due to the correlation constructs generated by the algorithms that can lead to biased results (Caliskan et al. 2017). The algorithms that can reproduce the traditional discrimination structures of members of disadvantaged groups is through the selection and weight given to the variables used by these systems for the measurement and prediction of the object they deal with. The priority given to some variables over others in measuring the phenomena that algorithms are responsible for can influence and bias the results (Soriano Arnánz 2021). Algorithmic discrimination can also be derived from errors or biases contained in the databases used in the development of automated decision-making systems. The existing literature has been concerned about the transfer and existence of gender biases dependent on decisions taken by the programmer or data scientist (Thelwall 2018; Scheuerman et al. 2019). In this sense, some of the experts interviewed indicate that the biggest problem is the perpetuation of certain roles and biases in society's consciousness. Therefore, in our opinion, more emphasis should be placed on the multidisciplinary education in AI, for raising the ethical and social awareness of AI professionals and

researchers, and for the education of society. Data quality is of particular importance for AI applications, dictating the need for mechanisms and metrics to safeguard this quality, especially to ensure that training data does not result in bias. Open access to public-sector data sources with high scientific and social value is one way of contributing to reducing such bias (Fernandez-Aller et al. 2021).

Regarding privacy problems, the results have shown concern about the impact of artificial intelligence on the protection of people's data. It is because public administrations have large databases of citizens that have been compiled in the exercise of their public functions. Although the processing of personal data through an algorithm is governed by the provisions of the GDPR, this European regulation may not be adapted to AI specialties (Wachter and Mittelstadt 2019) and therefore neither the Spanish regulation on privacy. However, some GDPR provision can be particularly useful since it enables a suitable design and a preventive analysis of the technological tools. For this reason, the public administrations that use AI algorithms must comply with a minimum set of conditions to guarantee the compliance of the data treatment. Furthermore, some experts interviewed emphasised that sometimes the systems that integrate the AI are in the form of third-party engines and components interlaced in the data processing of the data controller. These components can be libraries or source codes, but there can also be complete systems processing the data. In other words, there may be an information system in a public administration integrated into an AI engine that is running at the same time on a server in another country. This way of operating 'in the cloud' or with several chain managers is a source of risks that must be handled, especially if third countries intervene with inconsistent data protection legislation (Sobrino-García 2021b).

One of the characteristics of AI algorithms is the possibility of generating highly complex models, which would cause opacity according to the existing literature (Zlotnik 2019; Cerrillo i Martínez 2020c). In this regard, experts have exposed the different reasons for which algorithms may be opaque, ranging from technical conditions to legal causes. For this reason, the adoption of the principle of good administration is key to guaranteeing transparency due to the possibilities associated with good administration in its sense of efficacy and social precaution (Ponce Solé 2019; Cerrillo i Martínez 2020a). To grant the guarantee of the right to the good administration of citizens and allow the control of automated activity, a series of initiatives could be implemented. These would range from the construction of an administrative procedure for the algorithm's approval to the creation of an explanation right of the algorithm's operation. When public administrations make automated decisions, they must explain how the algorithm has worked and what kind of data it has used to obtain that result.

Finally, some interviewees identified that one of the conflicts derived from algorithm uses in the public sector was their legal status. Relative to this concern there is no agreement in the existing literature since some authors argue that algorithms are considered regulations (Boix Palop 2020). However, the Spanish courts have determined that the algorithm is public information if the algorithm is held by the administration. It is in line with the Italian and the Netherlands courts' judgements (2019–2020), which affirmed the responsibility to guarantee an algorithm cognoscible by public administrations. All this is has the aim of verifying that the decision adopted by the algorithm is under the prescriptions and purposes provided by the public administration rules.

5.3. The Lack of Spanish Regulation and Preventive Proposals

Finally, our principal results confirm the assumption on the third research question. Current Spanish legislation does not provide a clear answer to the risks of the use of AI by the public administration. For this reason, it is necessary to develop some mechanisms (legal and administrative) to respond to the problems associated with the use of algorithms in automated decision-making. In this regard, the existing literature recommends a strategy in which openness (open source, open data, open algorithms, etc.), can play an important role as an enabling factor, and in particular, facilitating dissemination and auditing

(Fernandez-Aller et al. 2021). However, according to the interviews performed, we believe that the approach of soft law norms is convenient (Sarmiento 2008) since the EU and Spain have a solid framework of principles with a human-centric approach. The regulatory base should be a standard with basic principles such as respect for human rights, dignity, freedom, democracy, and equality. Once this base is established, to mitigate problems with possible biases in the data, the creation of collegiate bodies or committees to control the development of AI projects is the most viable option, which is also accordant with the literature (Valero Torrijos 2020). These committees or collegiate bodies would have the participation of the people/groups that may be affected by the decision. Another proposal that can be assumed in the case of the public sector is the creation of certification systems such as ISO standards, according to which it would be verified that an entity complies with the ethical and complementary guidelines for the preparation of AI projects.

6. Conclusions

The expanding use of AI in public administration is triggering numerous opportunities for governments. Despite the benefits that this technology may involve, throughout this research it becomes clear that the use of artificial intelligence in public decision-making cannot be performed in any way since this can generate several difficulties. Governments cannot keep up with the rapid development of AI and the public administrations lack adequate AI policies. The starting point is the adoption of a legal definition of artificial intelligence since the EU has only defined this technology in non-binding documents. According to the analysis of the interviews and the previous literature, four key characteristics can be extracted for the definition of AI: autonomy, learning capacity, algorithm training, and the fact that it is a computer system. These elements can serve as the basis for the elaboration of a legal definition in laws or regulations and in accordance with the documents issued by European organisations. A legal concept of artificial intelligence is necessary to avoid insecurities at the legislative level.

The use of AI by the public sector can have a great impact, not only among governments but also among citizens. The organisational system and the traditionally bureaucratic form of government would be replaced by an automated and sometimes more efficient system. These challenges of AI use can be related to employees' lack of knowledge about artificial intelligence and machine learning, which would involve the incorporation of specialists and experts. According to the previous literature and the analysed interviews, the interpretation of AI can be complex, with a challenging opacity that could make it difficult to understand the system and communicate it to citizens. In this sense, AI could undermine the fundamental values of due process and transparency. Since these systems can consist of black-box processes, it is not always clear who is responsible for decisions made. On a social level, the use of AI could, in some cases, lead to the human workforce being replaced. However, people could be re-trained and re-employed for this new technology.

As has been demonstrated throughout the investigation, the use of AI by public administrations can entail several risks: opacity, legal uncertainty, biases, or breaches of personal data protection. However, the mechanisms already provided by Spanish law are not enough to avoid these risks as they have not been designed to face the use of artificial intelligence in public administration.

For this reason, we want to propose various measures that could be adopted for the regulation of AI according to the analysis carried out. At the internal level of the companies that are in charge of developing this technology, they could promote the adoption of ethical codes that take into account the gender issue through the the creation of groups of specialised professionals and the opening of code. In addition, another measure that could be implemented would be to carry out audits by independent third parties. The regulatory base should be a standard with basic principles such as respect for human rights, dignity, freedom, democracy, and equality. Once this base is established, to mitigate problems with possible biases in the data, the creation of collegiate bodies or committees to control the development of AI projects is the most viable option, which is also accordant with the

literature. Another proposal that can be assumed in the case of the public sector is the creation of certification systems such as ISO standards.

At the European level, a legal answer has been explored to help solve the risks of AI, following a series of principles centred on the human being. Consequently, at the end of April 2021, the Commission presented a proposal for a regulation on AI, “Proposal for a Regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and certain Union legislative acts”.

The research findings had limitations since this study provides an exploratory analysis of the Spanish perspective about the problems of the use of artificial intelligence among public administrations. Therefore, the findings might not be applicable to all the public administrations and governments of other countries. Future lines of research can advance the theoretical understanding of the impacts of AI on the public sector and the political perspective of the AI uses in public services, as well as the analysis and comparison of specific artificial intelligence projects in European countries.

In short, despite the limitations of this research, we developed an overview of the relevant risks derived from this new technology that Spanish public administrations may face and if the current law has mechanisms to deal with them. Additional research is needed to understand the specificities of the public administration AI uses in different countries of the EU, and how specific AI projects look like inside public administration to observe practical risks in the public sector. This will help identify how AI uses could differ from the different countries.

Funding: This research has received funding from the “Cátedra Feminismos 4.0 DEPO-UVIGO investigación” from 24 August 2020, from the project “Inteligencia Artificial e Xénero: as bases éticas e os problemas de sesgo de xénero na elaboración de algoritmos nas políticas públicas e procesos administrativos”, grant number 202000RS131H647. This publication reflects the views only of the author, and the Agency cannot be held responsible for any use, which may be made of the information contained therein.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The author declares no conflict of interest. The funder had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

- Abdeldayem, Marwam Mohamed, and Saeed Hameed Aldulaimi. 2020. Trends and Opportunities of Artificial Intelligence in Human Resource Management: Aspirations for Public Sector in Bahrain. *International Journal of Scientific & Technology* 9: 3867–71.
- Ali, Danish, and Samuel Frimpong. 2020. Artificial Intelligence, Machine Learning and Process Automation: Existing Knowledge Frontier and Way Forward for Mining Sector. *Artificial Intelligence Review* 53: 6025–42. [\[CrossRef\]](#)
- Allam, Zaheer, and Zaynah Asma Dhunny. 2019. On Big Data, Artificial Intelligence and Smart Cities. *Cities* 89: 80–91. [\[CrossRef\]](#)
- Ballester, Lluís, Carmen Orte, and Joseph Lluís Oliver. 2003. Análisis Cualitativo de Entrevistas. *Nómadas* 18: 140–49.
- Batarseh, Feras A., Laura Freeman, and Chih-Hao Huang. 2021. A Survey on Artificial Intelligence Assurance. *Journal of Big Data* 8: 1–30. [\[CrossRef\]](#)
- Bertolini, Massimo, Davide Mezzogori, Mattia Neroni, and Francesco Zammori. 2021. Machine Learning for Industrial Applications: A Comprehensive Literature Review. *Expert Systems with Applications* 175: 114–30. [\[CrossRef\]](#)
- Bogner, Alexander, Beate Littig, and Wolfgang Menz. 2009. Introduction: Expert Interviews—An Introduction to a New Methodological Debate. In *Interviewing Experts*. London: Palgrave Macmillan, pp. 1–13.
- Boix Palop, Andrés. 2020. Los Algoritmos Son Reglamentos: La Necesidad de Extender Las Garantías Propias de Las Normas Reglamentarias a Los Programas Empleados Por La Administración Para La Adopción de Decisiones. *Revista de Derecho Público* 1: 223–69.
- Caliskan, Aylin, Joanna J. Bryson, and Arvind Narayanan. 2017. Semantics Derived Automatically from Language Corpora Contain Human-Like Biases. *Science* 356: 183–86. [\[CrossRef\]](#)
- Campos Acuña, Ma Concepción. 2019. (In)Necesarias Regulaciones Para La Garantía Del Servicio Público. *Revista Vasca de Gestión de Personas y Organizaciones Públicas* 3: 74–91.
- Cerrillo i Martínez, Agustí. 2017. *Contratació Oberta*; Cataluña: Generalitat de Catalunya.

- Cerrillo i Martínez, Agustí. 2020a. ¿Son Fiables Las Decisiones de Las Administraciones Públicas Adoptadas Por Algoritmos? *European Law of Digital Administration & Law* 1: 18–36.
- Cerrillo i Martínez, Agustí. 2020b. El Derecho Para Una Inteligencia Artificial Centrada En El Ser Humano y Al Servicio de Las Instituciones. *Revista de Internet, Derecho y Política* 30: 1–6.
- Cerrillo i Martínez, Agustí. 2020c. El Impacto de La Inteligencia Artificial En Las Administraciones Públicas: Estado de La Cuestión y Una Agenda. In *Retos Jurídicos de La Inteligencia Artificial*. Edited by Agustí Cerrillo i Martínez and Miquel Peguera Poch. Pamplona: Thomson Reuters Aranzadi, pp. 75–92.
- Coglianesi, Cary, and David Lehr. 2017. Regulating by Robot: Administrative Decision Making in the Machine-Learning Era. *Georgetown Law Journal* 17: 1147–223.
- Cotino Hueso, Lorenzo. 2017. Big Data e Inteligencia Artificial. Una Aproximación a Su Tratamiento Jurídico Desde Los Derechos Fundamentales. *Dilemata* 24: 131–50.
- Cotino Hueso, Lorenzo. 2019. Riesgos e Impactos Del Big Data, La Inteligencia Artificial y La Robótica. Enfoques, Modelos y Principios de La Respuesta Del Derecho. *Revista General de Derecho Administrativo* 50: 1–30.
- Criado, J. Ignacio. 2016. Las Administraciones Públicas En La Era Del Gobierno Abierto. Gobernanza Inteligente Para Un Cambio de Paradigma En La Gestión Pública. *Revista de Estudios Políticos* 173: 245–75. [CrossRef]
- Criado, J. Ignacio. 2021. La Política de Administración Digital En España. De Los Servicios Públicos Digitales a La Gobernanza Inteligente y Administración Pública 4.0. In *Repensando la Administración Pública*. Edited by Carles Ramió. Madrid: Instituto Nacional de la Administración Pública, pp. 71–108.
- Dawes, Sharon S. 2013. Advancing Digital Government. The Research-Practice-Knowledge Connection. *Gestión y Política Pública*, 49–67.
- Dejonckheere, Melissa, and Lisa M. Vaughn. 2019. Semistructured interviewing in primary care research: A balance of relationship and rigour. *Family Medicine and Community Health* 7: 1–8. [CrossRef]
- European Commission. 2018. 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. Brussels. Available online: <https://digital-strategy.ec.europa.eu/en/library/communication-artificial-intelligence-europe> (accessed on 15 August 2021).
- European Parliament. 2019a. EU Guidelines on Ethics in Artificial Intelligence: Context and Implementation. Brussels. Available online: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI\(2019\)640163_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/640163/EPRS_BRI(2019)640163_EN.pdf) (accessed on 18 August 2021).
- European Parliament. 2019b. European Parliament Resolution of 12 February 2019 on a Comprehensive European Industrial Policy on Artificial Intelligence and Robotics. Strasbourg. Available online: https://www.europarl.europa.eu/doceo/document/TA-8-2019-0081_EN.html (accessed on 18 August 2021).
- European Parliamentary Research Service. 2019. *EU Guidelines on Ethics in Artificial Intelligence: Context and Implementation*. Brussels: European Parliament Research Service.
- Fernández Hernández, Carlos. 2021. Proyecto de Reglamento Europeo Sobre Inteligencia Artificial. *Diario La Ley. Ciberderecho* 50: 1–12.
- Fernandez-Aller, Celia, Arturo Fernandez de Velasco, Angeles Manjarres, David Pastor-Escuredo, Simon Pickin, Jesus Salgado Criado, and Txetxu Ausin. 2021. An Inclusive and Sustainable Artificial Intelligence Strategy for Europe Based on Human Rights. *IEEE Technology and Society Magazine* 40: 46–54. [CrossRef]
- Glaser, Barney G., and Anselm L. Strauss. 2000. *Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Routledge.
- Harris, John. 2018. Who Owns My Autonomous Vehicle? Ethics and Responsibility in Artificial and Human Intelligence. *Cambridge Quarterly of Healthcare Ethics* 27: 599–609. [CrossRef]
- Hildebrandt, Mirei. 2020. The Artificial Intelligence of European Union Law. *German Law Journal* 21: 74–79. [CrossRef]
- Huergo Lora, Alejandro. 2021. El Proyecto de Reglamento Sobre La Inteligencia Artificial. *Almacén de Derecho*. Available online: <https://almacendederecho.org/el-proyecto-de-reglamento-sobre-la-inteligencia-artificial> (accessed on 18 August 2021).
- King, Nigel, Christine Horrocks, and Joanna Brooks. 2019. *Interviews in Qualitative Research*. London: SAGE Publications.
- de la Sierra Morón, Susana. 2021. Control Judicial de Los Algoritmos: Robots, Administración y Estado de Derecho. *Revista de Jurisprudencia* 1: 1–12.
- López-Chao, Vicente, Antonio Amado Lorenzo, and Jorge Martín-Gutiérrez. 2019. Architectural Indoor Analysis: A Holistic Approach to Understand the Relation of Higher Education Classrooms and Academic Performance. *Sustainability* 11: 6558. [CrossRef]
- Martínez, Miguel. 2006. La Investigación Cualitativa (Síntesis Conceptual). *Revista de Investigación en Psicología* 9: 123–46. [CrossRef]
- Menéndez Sebastián, Eva M. 2021. *De la Función Consultiva Clásica a la Buena Administración. Evolución en el Estado Social y Democrático de Derecho*. Madrid: Marcial Pons.
- Mergel, Ines, Noella Edelman, and Nathalie Haug. 2019. Defining Digital Transformation: Results from Experts Interviews. *Government Information Quarterly* 39: 1–16. [CrossRef]
- Mesko, Bertalan. 2017. The Role of Artificial Intelligence in Precision Medicine. *Expert Review of Precision Medicine and Drug Development* 2: 239–41. [CrossRef]
- Miles, Matthew B. 1994. *An Expanded Sourcebook Qualitative Data Analysis*. London: SAGE Publications.
- Morley, Jessica, Luciano Floridi, Libby Kinsey, and Anat Elhalal. 2020. From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices. *Science and Engineering Ethics* 26: 2141–68. [CrossRef]

- Murphy, Kevin P. 2012. *Machine Learning: A Probabilistic Perspective*. Cambridge: The MIT Press.
- Navas Navarro, Susana. 2017. Derecho e IA Desde El Diseño. Aproximaciones. In *IA. Tecnología, Derecho*. Valencia: Tirant lo Blanch, pp. 23–72.
- Nikitas, Alexandros, Kalliopi Michalakopoulou, Eric Tchouamou Njoya, and Dimitris Karampatzakis. 2020. Artificial Intelligence, Transport and the Smart City: Definitions and Dimensions of a New Mobility Era. *Sustainability* 12: 2789. [\[CrossRef\]](#)
- Nishant, Rohit, Mike Kennedy, and Jacqueline Corbett. 2020. Artificial Intelligence for Sustainability: Challenges, Opportunities, and a Research Agenda. *International Journal of Information Management* 53: 102–4. [\[CrossRef\]](#)
- Noble, Safiya Umoja. 2018. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.
- Nowicki, Marek, Łukasz Górski, and Piotr Bała. 2021. PCJ Java Library as a Solution to Integrate HPC, Big Data and Artificial Intelligence Workloads. *Journal of Big Data* 8: 62–75. [\[CrossRef\]](#)
- O’Neil, Cathy. 2016. *Weapons of Math Destruction. How Big Data Increases Inequality and Threatens Democracy*. New York: Crown Publishers.
- Piñar Mañas, Jose Luis. 2019. Inteligencia Artificial: Técnica, Ética y Derecho. *Escritura Pública* 118: 40–41.
- Ponce Solé, Juli. 2018. La Prevención de Riesgos de Mala Administración y Corrupción, La Inteligencia Artificial y El Derecho a Una Buena Administración. *Revista Internacional de Transparencia e Integridad* 6: 1–19.
- Ponce Solé, Juli. 2019. Inteligencia Artificial, Derecho Administrativo y Reserva de Humanidad: Algoritmos y Procedimiento Administrativo Debido Tecnológico. *Revista General de Derecho Administrativo* 50: 1–30.
- Qu, Sandy Q., and John Dumay. 2011. The Qualitative Research Interview. *Qualitative Research in Accounting & Management* 8: 238–64.
- Ragin, Charles C. 2009. Reflections on Casing and Case-Oriented Research. In *The SAGE Handbook of Case-Based Methods*. London: SAGE Publications, pp. 522–50.
- Raso, Filippo, Hannah Hiilligoss, Vivek Krishnamurthy, Christopher Bavitz, and Levin Yerin Kim. 2018. Artificial Intelligence & Human Rights: Opportunities & Risks. *SSRN Electronic Journal* 1: 1–63.
- Sarmiento, Daniel. 2008. *El Soft Law Administrativo. Un Estudio de Los Efectos Jurídicos de Las Normas No Vinculantes de la Administración*. Madrid: Thomsom-Civitas.
- Scheuerman, Morgan Klaus, Jacob M. Paul, and Jed R. Brubaker. 2019. How Computers See Gender. *Proceedings of the ACM on Human-Computer Interaction* 3: 1–33. [\[CrossRef\]](#)
- Silverman, David. 2013. *Interpreting Qualitative Data*. London: SAGE Publications.
- Sobrinó-García, Itziar. 2021a. Desafíos y Limitaciones En La Contratación Pública: El Impacto de La Protección de Datos Tras Los Últimos Cambios Legislativos. *Revista General de Derecho Administrativo* 56: 1–30.
- Sobrinó-García, Itziar. 2021b. Las Decisiones de Adecuación En Las Transferencias Internacionales de Datos. El Caso Del Flujo de Datos Entre La Unión Europea y Estados Unidos. *Revista de Derecho Comunitario Europeo* 25: 227–56. [\[CrossRef\]](#)
- Soriano Arnánz, Alba. 2021. Decisiones Automatizadas y Discriminación: Aproximación y Propuestas Generales. *Revista General de Derecho Administrativo* 56: 1–30.
- Sun, Tara Qian, and Rony Medaglia. 2019. Mapping the Challenges of Artificial Intelligence in the Public Sector: Evidence from Public Healthcare. *Government Information Quarterly* 36: 368–83. [\[CrossRef\]](#)
- Tegmark, Max. 2017. *Life 3.0: Being Human in the Age of Artificial Intelligence*. New York: Penguin Books.
- Thelwall, Mike. 2018. Gender Bias in Machine Learning for Sentiment Analysis. *Online Information Review* 42: 343–54. [\[CrossRef\]](#)
- Todolí Signes, Adrián. 2020. Retos Legales En El Uso Del Big Data En La Selección de Sujetos a Investigar Por La Inspección de Trabajo y de La Seguridad Social. *Revista Galega de Administración Pública* 59: 313–37. [\[CrossRef\]](#)
- Valcárcel Fernández, Patricia. 2019. La Colaboración Público-Privada En La LCSP. Algunas Experiencias En El Ámbito Local. In *La Gestión de Los Servicios Públicos Locales En El Marco de La LCSP, La LRJSP y La LRSAL*. Edited by J. M. Martínez Fernández. Madrid: Wolters Kluwer.
- Valero Torrijos, Julián. 2020. The Legal Guarantees of Artificial Intelligence in Administrative Activity: Reflections and Contributions from the Viewpoint of Spanish Administrative Law and Good Administration Requirements. *European Review of Digital Administration & Law* 1: 55–61.
- Valle-Cruz, David, Edgar Alejandro Ruvalcaba-Gomez, Rodrigo Sandoval-Almazan, and J. Ignacio Criado. 2019. A Review of Artificial Intelligence in Government and Its Potential from a Public Policy Perspective. Paper Presented at 20th Annual International Conference on Digital Government Research, Dubai, United Arab Emirates, June 18–20; pp. 91–99.
- Valle-Cruz, David, J. Ignacio Criado, Rodrigo Sandoval-Almazán, and Edgar A. Ruvalcaba-Gomez. 2020. Assessing the Public Policy-Cycle Framework in the Age of Artificial Intelligence: From Agenda-Setting to Policy Evaluation. *Government Information Quarterly* 37: 1–15. [\[CrossRef\]](#)
- Valles, Miguel S. 2009. *Entrevistas Cualitativas*. Madrid: Centro de Investigaciones Sociológicas.
- Van Roy, Vicent, Fiammetta Rossetti, Karine Perset, and Laura Galindo-Romero. 2021. *AI Watch—National Strategies on Artificial Intelligence: A European Perspective*, 2021th ed. JRC Working Papers No. JRC122684. Brussels: Publications Office of the European Union.
- Vesnic-Alujevic, Lucia, Susana Nascimento, and Alexandre Pólvora. 2020. Societal and Ethical Impacts of Artificial Intelligence: Critical Notes on European Policy Frameworks. *Telecommunications Policy* 44: 1–15. [\[CrossRef\]](#)
- Wachter, Sandra, and Brent Mittelstadt. 2019. A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI. *Columbia Business Law Review* 2: 1–130.

-
- Wirtz, Bernd W., Jan C. Weyerer, and Carolin Geyer. 2019. Artificial Intelligence and the Public Sector—Applications and Challenges. *International Journal of Public Administration* 42: 596–615. [[CrossRef](#)]
- Zheng, Yongqing, Han Yu, Yuliang Shi, Kun Zhang, Shuai Zhen, Lizhen Cui, Cyril Leung, and Chunyan Miao. 2020. PIDS: An Intelligent Electric Power Management Platform. *Proceedings of the AAAI Conference on Artificial Intelligence* 34: 13220–27. [[CrossRef](#)]
- Zlotnik, Alexander. 2019. Inteligencia Artificial En Las Administraciones Públicas: Definiciones, Evaluación de Viabilidad de Proyectos y Áreas de Aplicación. *Boletín* 84: 25–32.