Government Information Quarterly 39 (2022) 101679

Contents lists available at ScienceDirect

Government Information Quarterly

journal homepage: www.elsevier.com/locate/govinf

Assessing behavioral data science privacy issues in government artificial intelligence deployment

Jose Ramon Saura [a,*], Domingo Ribeiro-Soriano [b], Daniel Palacios-Marqués [c]

[a] Rey Juan Carlos University, Madrid, Spain
[b] Universitat de Valencia, Valencia, Spain
[c] Universitat Politècnica de València, Valencia, Spain

ARTICLE INFO

Keywords:
Behavioral data sciences
Governments
Collective behavior analysis
Artificial intelligence
Surveillance capitalism
Privacy

ABSTRACT

In today's global culture where the Internet has established itself as the main tool for communication and commerce, the capability to massively analyze and predict citizens' behavior has become a priority for governments in terms of collective intelligence and security. At the same time, in the context of novel possibilities that artificial intelligence (AI) brings to governments in terms of understanding and developing collective behavior analysis, important concerns related to citizens' privacy have emerged. In order to identify the main uses that governments make of AI and to define citizens' concerns about their privacy, in the present study, we undertook a systematic review of the literature, conducted in-depth interviews, and applied data-mining techniques. Based on our results, we classified and discussed the risks to citizens' privacy according to the types of AI strategies used by governments that may affect collective behavior and cause massive behavior modification. Our results revealed 11 uses of AI strategies used by the government to improve their interaction with citizens, organizations in cities, services provided by public institutions or the economy, among other areas. In relation to citizens' privacy when AI is used by governments, we identified 8 topics related to human behavior predictions, intelligence decision making, decision automation, digital surveillance, data privacy law and regulation, and the risk of behavior modification. The paper concludes with a discussion of the development of regulations focused on the ethical design of citizen data collection, where implications for governments are presented aimed at regulating security, ethics, and data privacy. Additionally, we propose a research agenda composed by 16 research questions to be investigated in further research.

1. Introduction

In recent years, the development of artificial intelligence (AI) has led to the adaptation of organizational models in both companies and public organizations (Brynjolfsson & Mitchell, 2017). In today's global culture where the Internet has established itself as the main tool of communication, the global system of economy and regulations, as well as data and decisions based on behavioral analysis, have become essential for public actors (Ballestar, Camiña, Díaz-Chao, & Torrent-Sellens, 2021; Irvin & Stansbury, 2004).

In the context of this connected society, conceptualization, definition, and establishment of both theoretical and legal parameters that would set ethical and efficient limits on the analysis, treatment, and use of citizens' data have become a challenge in scientific, legal, and professional settings (Kamolov & Teteryatnikov, 2021; Narayanan, Huey, & Felten, 2016). As studied by Zuboff (2019b) numerous concerns regarding user privacy have emerged—particularly, when setting parameters for governments to make decisions regarding how to apply AI to understand behaviors in the society (Hiller & Bélanger, 2001), predict its actions and movements (Altman, Wood, O'Brien, Vadhan, & Gasser, 2015), and act accordingly. Of note, AI refers to the simulation of human intelligence linked to the development of algorithmic models that automatically work and learn by themselves through inputs developed by humans (Nagtegaal, 2021).

Abbreviations: AI, Artificial Intelligence; BDS, Behavioral Data Sciences; UGC, User-Generated Content; UGD, User-Generated Data; IoT, Internet of Things; MB, Machine Behavior; AB, Algorithmic Behavior; BA, Behavioral Analytics; BE, Behavioral Economics; CBA, Collective Behavior Analysis; NLP, Natural Language Processing; LDA, Latent Dirichlet Allocation; TA, Textual Analysis.

* Corresponding author.
E-mail addresses: joseramon.saura@urjc.es (J.R. Saura), domingo.ribeiro@uv.es (D. Ribeiro-Soriano), dapamar@doe.upv.es (D. Palacios-Marqués).

https://doi.org/10.1016/j.giq.2022.101679
Received 21 January 2021; Received in revised form 31 January 2022; Accepted 1 February 2022
Available online 21 February 2022

In recent years, the use of AI by corporations and governments has grown exponentially (Zuiderwijk, Chen, & Salem, 2021), and this growth has been predetermined by many benefits of AI, such as analysis of large amounts of data, predictions with high accuracy rates, identification of trends and patterns, predictions of complex associations, improvement of profitability, analysis of financial ratios and risks, among others uses (Mikalef et al., 2021).

In parallel to the increase in the use of AI in governments, and as a consequence of the evolution in data and behavioral analysis practices, the concept of behavioral data sciences (BDS) has been developed to combine a multitude of issues related to data science and behavior (Harari et al., 2016). Although the very term "behavioral data sciences" does not appear in the scientific literature, several previous studies, including Agarwal and Dhar (2014) and Van Der Aalst (2016), have directly defined the future guidelines for its development. Therefore, the term BDS refers to a new and emerging interdisciplinary field that combines techniques from behavioral sciences, psychology, sociology, economics, and business, and uses the processes from computer science, data-centric engineering, statistical models, information science, or mathematics, in order to understand and predict human behavior using AI (Saura, Palacios-Marqués, & Iturricha-Fernández, 2021). In essence, BDS is a mix of disciplines that combines knowledge of the data that users or citizens publicly generate on the Internet -known as user-generated content (UGC) or Data (UGD)- through the use of mobile applications and other connected devices, such as Internet of Things (IoT), smart homes, self-driven cars, or through smart-cities connected services (Schreiner, Fischer, & Riedl, 2019).

With the use of techniques focused on BDS, governments could apply algorithms that work with AI and systems that analyze behavior (Grimmelikhuijsen, Jilke, Olsen, & Tummers, 2017), identify patterns to explore the knowledge about the society (Men & Tsai, 2014) as well as its consumers or users (Chen et al., 2020; Chen et al., 2021). In this study, the use of the BDS concept is specifically linked to the analysis of AI strategies developed by governments to date. Since the term does not relevantly appear in the published scientific literature, the present study is pioneering and original in this respect.

Furthermore, corporates not only leverage users and clients' data to improve their products and services, but also use them as exchange currency with interested third parties, such as governments or other public institutions (Silverman, 2017). Therefore, by studying user behavior data, companies and governments develop sophisticated power machines that predict an economic logic that helps corporates generate more money at the expense of users and citizens (Zuboff, 2019a). Likewise, according to the government actions, the use of AI raises concerns about privacy and personal security issues (Yang, Elisa, & Eliot, 2019). While predictions are not equal to observations, the more data is obtained from the society, the greater is the ability to predict. Accordingly, predictions can reach the same level of effectiveness as that of observations (Zuboff, 2019a). Therefore, if governments have this intelligence, and if it is also automated based on AI, the risk to privacy and free decision-making in the society could be at threat (Mazurek & Małagocka, 2019).

In this context, a key notion in this field is the concept of surveillance capitalism. According to Cinnamon (2017) and Zuboff (2019b), in surveillance capitalism, user experience and behavioral data are used as economic drivers to create a new economy where economic drivers and profits come from predicting how users behave. Therefore, considering this new concept, governments can take action and use AI as a tool focused on BDS. However, as stated Bromberg, Charbonneau, and Smith (2020), such use can violate citizens' privacy and security. For example, by using AI and BDS, governments can interfere with the behavior of the society to achieve a change in behavior, without the society being aware of it (Zuboff, 2015). There is also evidence of how governments can use AI to predict election results using massive data to change the voting intentions of thousands of users (Isaak & Hanna, 2018). This was the case of US Facebook users' behavioral data that, when analyzed with behavioral prediction algorithms, such as the one developed by Cambridge Analytica (Heawood, 2018), were employed to modify the election results in the US presidential campaign between Donald Trump and Hilary Clinton in 2016 (Cadwalladr & Graham-Harrison, 2018). Another example could be the famous German doll, Cayla, which recorded chunk dialogues said by children (Haynes, Ramirez, Hayajneh, & Bhuiyan, 2017), the company then sold those data Nuance Communications, which, in turn, developed a voice recognition software and sold it to the US Central Intelligence Agency (CIA) (Madnick, Johnson, & Huang, 2019). In this case, we can speak about government suppliers' provision of AI-related services that are unethical from citizens' point of view.

In this context, after the development of this type of events where AI, governments, and the data collection capacity of corporations is questioned, essential questions regarding the knowledge, authority, and power of government use of BDS techniques should be explored. Furthermore, understanding the predictive ability that government institutions might obtain if they train AI models that can predict user behavior is a prerequisite for any society to feel confident about implementing new technologies (Hobolt, Tilley, & Wittrock, 2013). Of note, data predictions and models that work with the prediction of human behavior are becoming dominant forms of capitalism and generate new business models and new products in the form of data (Zuboff, 2019b). Of note, BDS is a clear priority for the development of ethical strategies by governments when they implement AI in their strategies as it is presented as a new concept linked to user privacy, AI deployment in governments, or behavioral analytics, that brings together all of the above in the form of analysis of society's behavioral data.

However, several unanswered questions remain, such as what is the legitimacy of predicting user behavior? And who do these behavioral data belong to? Based on the privacy concerns outlined above and the originality of the study justified under the BDS new emerging concept, to the best of our knowledge, none of previous studies had identified and described the risks of governmental implementation of AI to citizens' privacy. Furthermore, there has been no research linking the concept of BDS to the main uses of AI by governments. Thus, we seek to fill a gap in the literature by exploring the possible uses and risks to citizens' privacy if governments implementation of AI in their strategies under the new BDS conceptual framework. To this end, this study first develops a systematic review of the literature to establish and confirm the main scientific contributions to date in this field of study. Secondly, based on the results of the systematic review of the literature, 15 interviews were conducted with 11 individuals working in the government; of these, 2 were economists for the government, and 2 belonged to organizations that advise the government. Thirdly, based on the coded results of the interviews, two data-mining techniques (topic-modeling and textual analysis) were developed to identify insights and create knowledge related to the object of study. Following this approach, the present study aims to identify and discuss the main practical and theoretical implications for governments when using AI-based strategies with BDS techniques.

Therefore, in order to cover the identified gap in the literature, the present study addresses the following research questions (RQ): RQ1: What kind of citizens' privacy issues are expected when governments use behavioral-based AI in their strategies? and RQ2: What AI techniques can governments develop to predict the society's behavior?

With the development of the study and the answers of the RQ, this study also intends to attain the following specific objectives:

- To identify definitional perspectives of behavioral data science privacy issues in government AI deployment
- To explore the types of behavioral data science approaches used in governments
- To create knowledge about government AI deployment preserving society privacy

- To outline future guidelines to track new challenges in behavioral analytics and government AI deployment

Based on the results, we discuss theoretical implications regarding the application of AI strategies used by governments that respect the privacy of citizens' data. In addition, the main contributions to date are theorized in relation to the management of user data and the need to regulate security, ethics, and privacy of user data. Similarly, we also discuss practical implications that form a guide for the application of AI strategies by governments that avoid any type of privacy violations linked to surveillance capitalism actions.

The remainder of this paper is structured as follows. In Section 2, the theoretical framework of the study is presented. Section 3 discusses the methodological approaches used. Section 4 reports the results. Section 5 provides a discussion of important theoretical contributions and future directions that our results offer for the analysis of BDS privacy issues in government AI deployment. Conclusions, along with a discussion of theoretical and practical implications, are presented in Section 6.

2. Theoretical framework

2.1. Understanding surveillance capitalism and behavioral data sciences

As argued by Zuboff (2019b) and Belhadi et al. (2021), we are living in one of the deepest transitions in the information age—namely, in an ecosystem where data are the largest source of information. Seeking to outline a theoretical background with the main concepts used to analyze and predict user behavior in the digital ecosystem, this section identifies the main theoretical perspectives used in the literature to analyze the factors that contribute to the development of AI in governments.

For their part, governments need to be updated and use the latest technologies to understand what the demands of the society are (Figenschou, 2020). However, according to many initiatives, the regulation of the Internet itself is not working, and the society demands that its data should remain anonymous at all costs (Zuboff, 2019b). This raises concerns about user privacy (Ribeiro-Navarrete, Saura, & Palacios-Marqués, 2021). Users are aware of the fact that, based on the analysis of human experiences linked to behavioral data, governments can turn their actions into sophisticated intelligent machines capable of predicting any issue targeted by governments (Kavanaugh et al., 2012). Therefore, the ultimate goal will always be to understand the future behavior of the society regulated by governments (Linders, 2012).

Under this paradigm of privacy concerns about AI and its implementation by governments to monitor, actively listen, trace possible states of alarm, or predict any kind of event that negatively affects the society, the concept of surveillance capitalism is born (Cinnamon, 2017; Zuboff, 2015; Zuboff, 2019b). The concept of surveillance capitalism advocates that human experience is unilaterally automated as data sources to predict human behavior (Zuboff, 2019a). While there are indeed objectives of service improvement and understanding the society's behavior to improve the public offer by governments (Andrew & Baker, 2019), the concept of surveillance capitalism also implies that humans are used as products of massive data production to improve the economic profitability of companies at the expense of the data about user behavior (Zuboff, 2015).

In the circumstances where ethical actions are lacking, companies and governments use behavioral data to make the society behave in ways that are more convenient to obtain greater economic benefits (Zuboff, 2015). When viewed from a business perspective, this leads to an increasing number of Internet-centered business models that cater to addictive behavioral patterns (Hou, Xiong, Jiang, Song, & Wang, 2019). In this way, users generate more data about their behavior; accordingly, their attitudes and feelings can be predicted. Then, based on these actions, companies and governments generate more profitability on the advertising (Palos-Sanchez, Saura, & Martin-Velicia, 2019) products shared in these business models (Dwivedi, Kapoor, & Chen, 2015), or using user behavior data as the basis of data-centered strategies (Dwivedi et al., 2018).

In surveillance capitalism, the main source of data is the information generated by users while using connected devices. All this information is analyzed using BDS that takes a new perspective of analysis through a combination of different fields of research (Zhuoxuan, Yan, & Xiaoming, 2015). In recent years, the number of tools used by both governments and companies to obtain data has considerably increased (Paul & Aithal, 2020). In fact, many variables indicate parameters for measuring user behavior on the Internet or through their mobile and connected devices (Hobolt et al., 2013).

Until now, the main sources of data were websites, cell phones, intelligent organization systems, Customer Relationship Management (CRM) systems, and marketing automation sources, among others. This type of data always generates categories known as events or objectives, which have the purpose of explaining some properties defined by the organizational structure of the data-analysis system (Abou Elassad, Mousannif, Al Moatassime and Karkouch, 2020). However, as mentioned above, the number and type of connected devices has recently exponentially increased, from IoT to smart city services, among other connected devices (Kankanhalli, Charalabidis, & Mellouli, 2019).

The understanding of user behavior data on the Internet has led to the emergence of new digital marketing strategies in the business ecosystem (Dwivedi et al., 2020). It is not the first time that the business ecosystem offers opportunities and benefits to government institutions to maximize their processes (Zhang, Wang, & Zhu, 2020), increase the efficiency of their strategies (Pencheva, Esteve, & Mikhaylov, 2020), or create new listening tactics (Macnamara, 2015). Following these considerations, Table 1 presents the main concepts related to BDS analysis that can be used by governments to monitor user behavior through the data they generate.

Table 1
Main concepts linked to the analysis of behavioral data.

Concept	Description	Authors
Behavioral Data Sciences (BDS)	An interdisciplinary field that studies user behavior through the data they generate from sociological, psychological, and economic perspectives applying statistics, mathematics, and data automation.	Litman, Robinson, and Abberbock (2017) Xu, de Barbaro, Abney, and Cox (2020)
Machine Behavior (MB)	A field that leverages behavioral sciences to understand the behavior of AI agents.	Abou Elassad, Mousannif, Al Moatassime and Karkouch (2020) Oey, Jones, Bullard, and Sant (2020)
Algorithmic Behavior (AB)	A field of study of the BDS using algorithms in large databases	Hobolt et al. (2013) Macnamara (2015)
Behavioral Analytics (BA)	Study of user behavior data using the Internet and social networks	Touma, Bertino, Rivera, Verma, and Calo (2017) Khan (2017)
Behavioral Economics (BE)	Studies the effect of the cognitive and emotional psychology of culture and society on the predictions of economic theory	Hursh (1984) Streletskaya et al. (2020)
Big Behavioral Data Science (BBDS)	Refers to BDS and the use of Big Data techniques	Gomez-Marin, Paton, Kampff, Costa, and Mainen (2014)
Behavior Informatics (BI)	Investigates user behavior data with processes focused on computer sciences.	Cao et al. (2014) Paul and Aithal (2020)
Collective Behavior Analysis (CBA)	Various approaches to the study of behavior focused on the collective knowledge of individual behavior	Belhadi et al. (2021) Turner and Killian (1957)
Behavior Learning Analysis (BLA)	A scientific field for the study of behavior with data entry learning techniques	Zentall, Galizio, and Critchfield (2002) Zhuoxuan et al. (2015)

Source: The authors.

2.2. Main user's behavior data sources used by the government in their monitoring strategies

In the ecosystem that drives the development of the economy based on behavioral analysis and its data, the importance of data sources can hardly be overestimated (White & Boatwright, 2020). As argued in many previous studies, since users are not fully aware that their data will be used and sold to interested third parties for a financial contribution (Acar, Englehardt, & Narayanan, 2020), ethics is not an essential component of new business models focused on massive data collection and analysis (Löfgren & Webster, 2020).

Moreover, several available initiatives—such as the new GDPR legislation introduced by the European Union through the European Commission to protect users based on abusive uses of their data—are insufficient (Sørensen & Kosta, 2019). Although the new regulation obliges companies to explicitly specify how the data are used, in reality, users do not possess knowledge necessary to understand privacy policies and legal notices of the applications on their mobile computers and any types of connected device (Martin, 2015).

There is evidence that, due to the psychological phenomenon known as "instantaneous reward", despite some awareness about privacy issues, the millennial generation and the digital natives prefer to use the applications as fast as possible instead of taking time to understand how their data will be used (Hull et al., 2004). In many situations, including the recent Covid-19 pandemic, the issue of user privacy has raised many concerns (Maher, Hoang, & Hindery, 2020). During the Covid-19 crisis, in order to track Covid-19 infections and notify citizens if they have been in contact with an infected, many governments have decided to ask citizens to use applications that track their location (Gerard, Imbert, & Orkin, 2020).

A recent analysis of these novel monitoring techniques by governments suggested that citizens frequently use this type of active listening (Maher et al., 2020; Zhou, Yang, Xiao, & Chen, 2020). The sources of data that governments may have access have been studied previously within several projects, such as the one published by *The New York Times* (Thompson & Warzel, 2019). Specifically, Thompson and Warzel (2019) highlighted many decision-making concerns that governments may have with access to multiple companies collecting information from users. Then, governments use user data to improve their processes of monitoring the society and its behavior (Thompson & Warzel, 2019), thereby prioritizing the issue of national security. With this type of strategy, user behavioral data are used as a source that governments use to train their algorithms that work with machine learning. Accordingly, behavioral data analysis is of a paramount important for governmental strategies focused on AI (Ribeiro-Navarrete et al., 2021).

Of note, user data can be transferred by third parties to governments (Thompson & Warzel, 2019). As described by Saura, Ribeiro-Soriano, & Palacios-Marqués (2021b), data sources on user behavior can be of the following three types: (i) public, i.e. when users are aware that the information they generate is in the public domain; (ii) private, i.e. when users know that the information they generate will be used exclusively for their personal use, and (iii) when behavioral data are transferred to third parties as products, including governments, public or private institutions (Saura, Ribeiro-Soriano, & Palacios-Marqués, 2021c). Table 2 summarizes data sources that can be used by governments for the BDS analysis.

The data sources shown in Table 2 are examples of the multitude of citizen behavior data sources that can be used by governments to obtain information for further analysis with AI (Cate, 2008). In this context, it is unsurprising that the society's concerns about privacy continue to grow (LaBrie, Steinke, Li, & Cazier, 2018). If used by governments in their systems for control, prediction, and analysis of user behavior, these data sources can affect the privacy and security of citizens' personal data.

Table 2
Data sources that can be used by the government to deploy AI strategies.

Data sources	Description	Public access	Private access	Third parties
User-generated data (UGD)	The data publicly generated by users in digital ecosystems	√		√
User-generated content (UGC)	The content published by users of social networks and online platforms	√		
User-generated behavior (UGB)	The set of connotations derived from user online behavior		√	√
Internet history	History of user searches and website visits		√	
Digital customer journey	Map of user actions to make a purchase, visit a website, or send information		√	√
User location	One of the fundamental indicators to measure the movement of the society. It can be consulted through its intelligent devices, such as smartphones or smartwatches.		√	√
Connected devices	Connected devices such as thermostats, home assistants, lamps, bulbs, etc.		√	√

Source: The authors.

3. Methods

3.1. Systematic review of the literature

To better understand the main uses of AI by governments as studied in the scientific literature to date, we conducted a systematic review of the literature (de Camargo Fiorini, Seles, Jabbour, Mariano, & de Sousa Jabbour, 2018). Systematic literature reviews are exploratory research approaches used to understand emerging new fields of study (Kraus, Breier, & Dasí-Rodríguez, 2020). A major reason underlying the recent increase in the number of systematic literature reviews is that a literature review makes it possible to outline a theoretical framework with the main agents that contribute to the development of the proposed research objective. Therefore, the aim of systematic literature review is to analyze an emerging issue and to identify the main techniques employed to study that issue. Therefore, systematic reviews are an effective method to identify the proposed objectives related to AI uses in governments and citizens privacy (Zuiderwijk et al., 2021).

In the present study, we followed the procedure developed by Bem (1995), who proposed that a systematic review should be divided into the following three steps. In the first step, the topics to be discussed within the scientific area are identified. To this end, keywords are identified that can summarize the objective of the research through searching databases (Sarkis, Zhu, & Lai, 2011). In the second step, the searches in these databases are performed, the collected data are filtered, and the results are analyzed (Akter & Wamba, 2016). During the filtering process, titles, abstracts, and keywords of potentially relevant studies are examined. This is followed by the analysis of the content of the articles, and their suitability for the review is assessed. The studies that do not meet these criteria are excluded from the systematic review process. In the third step, the content of the contributions retained in the sample is analyzed, and the main concepts are discussed (Zeng, Hu, Balezentis, & Streimikiene, 2020).

In the present study, final contributions were selected during the review process that focused on identification of the main purposes of each potentially relevant study (Akter et al., 2019). The searches were conducted in the following databases: Web of Sciences (WOS), IEEE Xplore, ScienceDirect, ACM Digital Library, and AIS Electronic Library. The keywords used to search the databases were "Government" OR "Governance" OR "Public Management" OR "Public Sector" OR "Public Administration" OR "Public Policy" OR "State" OR "Municipality" OR

"Citizens" AND "Artificial Intelligence" OR "AI" OR "Predictive Analytics" OR "Intelligence Systems" OR "Expert Systems" OR "Collective Behavior" OR "Surveillance Capitalism" OR "Behavioral Analysis". The searches were performed between October 5 and 10, 2020 and updated in January 2022. Of note, the search term BDS has not been used in this process, since the results of government studies using AI were analyzed from the perspective of BDS as a new emerging concept, which this study thoroughly outlines and defines in the results.

The results of the process were as follows. In WOS, 20 articles were selected from a total of 65 potentially relevant results; in ScienceDirect 7 results were selected from a total of 29 studies; in AIS Library, the total number of potentially relevant studies was 3, of which only 1 was retained in the final dataset; ACM Digital Library, a total of 4 studies were found, of which 2 were selected; finally, in IEEE Xplore, of a total of 20 potentially relevant study, 4 were selected. Therefore, after the selection process, a total of 34 research studies were selected to be included in the present study. For the exclusion criteria, we followed PRISMA evidence-based minimum set of items (Saura, Ribeiro-Soriano, & Palacios-Marqués, 2021a) aimed to filter quality research studies. First, the abstracts and keywords of the articles were analyzed to identify inadequate and not inclusive terms related to the objectives of the study. Second, an in-depth analysis of the articles identified as suitable was performed. Next, we analyzed whether the objectives of the study were directly or indirectly linked to the objectives of the present research. Then, we determined whether the topic is related to the research objectives. Additionally, we identified whether or not the quality of the methodology and evaluation of results were acceptable. Finally, articles that did not describe or specify terms appropriately to the objectives of the present study were excluded. Accordingly, Table 3 provides further detail on the studies included in the present study that were analyzed.

Following the methodological indications outlined in Snelson (2016) and Collins et al. (2021) and Mikalef et al. (2021), once the systematic review of the literature was developed to verify the validity of theoretical underpinnings, we proceeded to the development of the second part of the proposed approach. In this way, once the relevancy of AI and BDS governments uses was justified, we structured and designed the interviews based on the results of the systematic literature review. Details of this approach are presented below.

3.2. In-depth interviews

Seeking to obtain additional knowledge regarding the uses of AI by governments and the concerns related to citizens' privacy, we conducted in-depth interviews with informants working in governments. Following the guidelines proposed by MacDougall and Fudge (2001), our qualitative interviews were held with politicians, senators, and other government-related officials in Spain.

The ultimate goals of these interviews was not to quantitatively assess the studied phenomenon, but rather to gain a deep understanding of it by obtaining information from an original primary source. The importance of such qualitative approach was previously justified by Orlikowski and Baroudi (1991) and Roberts (2015). Subsequently, the content of the interviews was used to build theory and extract insights.

We conducted a total of 15 interviews on user privacy and AI strategies developed by governments. Of these 15 interviews, 5 were conducted by phone (Pell et al., 2020), 2 by video call (Lukacik, Bourdage, & Roulin, 2020), 4 in person (Lukacik et al., 2020), and 4 by email (McKinley, Fong, Udelsman, & Rickert, 2020). In all cases, the interviews were digitally coded for further analysis under the Natural Language Processing (NLP) framework. Of 15 informants, 11 individuals worked in the government, 2 were economists for the government, and 2 belonged to organizations that advise the government (see Table 7). The informants were Spanish (12), Venezuelan (1), Egyptian (1), and Colombian (1) nationals. Their identities are anonymized in the present study (Natow, 2020). The three interviewees who were not native Spaniards live in Spain and work for governments or corporations linked to economics, finance, and politics. All members taking part in the interviews are linked to the Club Financiero Génova (CFG) in Madrid, a club focused on economic development, business, and politics. The interviewees were informed of the interviews at various events held at the CFG and contacted afterwards. The interviews were conducted in Spanish and translated into English.

Of note, as informed by the European Commission, Spain has developed a strategy report to monitor the development, as well as to uptake and measure the impact of AI in their government actions. The Spanish government has informed that they use AI to facilitate the development and deployment of the economy and society. Its strategy adopts a multidisciplinary approach to address economic, social, environmental, public management, and governance challenges, and it includes perspectives for a wide range of sectors and disciplines (European Commission, 2020).

In-person interviews and video call interviews lasted for about 30–40 min each. Telephone interviews averaged 20–25 min in length. Email interview responses averaged 750–600 words each. Interview data were collected between October 15, 2020, and January 8, 2021. Questions are shown in Appendix A. The informants were selected based on the work they do or have previously done in the government. All informants were linked to public administrations, governments, political parties, or advisors to the government. Our interviews were semi-structured and included open-ended questions. Table 4 shows the characteristics of our informants based on their role, industry of specialization, professional status, organization they belong to, and nationality.

The main reason to ask open-ended questions in our interviews was to address a wider range of experiences (Dhillon & Torkzadeh, 2006). As noted above, the interview data were then transcribed and coded using exploratory data-based techniques (Bacq, Janssen, & Noël, 2019; Cooke-Davies & Arzymanow, 2003). The interviews received via email were used directly in the original format and sent for coding in the global database. The demographic characteristics of the informants are summarized in Table 5.

3.3. Data-mining techniques: Using LDA and TA to extract insights

In the last decade, data-mining techniques have been extensively used notably in the scientific literature (Yang & Wu, 2006). These techniques are used to create knowledge and extract insights from both structured and unstructured databases (Wu et al., 2003). A combination of several data-mining techniques processes can provide truly relevant insights into the objects proposed under study (Jindal & Borah, 2013).

In the present study, two data-mining processes were combined: Latent Dirichlet Allocation (LDA) and Textual Analysis (TA). The first one was a topic-modeling algorithm developed in Python to extract insights in the form of topics. LDA was applied to the database containing the content of the in-depth interviews (Blei, Ng, Jordan, & Lafferty, 2003; Pritchard, Stephens, & Donnelly, 2000). The novelty of this approach is that we used a methodology typically applied to analyze to explore primary interview data. These considerations are indicated in Krippendorff (2013) for the process of content analysis.

Specifically, the algorithm applied by the LDA identifies the most relevant words in the analyzed documents. In the present study, each interview was considered as a document. Using the topic-modeling process with LDA, we identified approximately 10 words for each document. These words were then used to form the names of topics in the data. This is a standard process in the use and development of LDA using the NLP framework. In the present study, the LDA process was computed with Python LDA 1.0.5 software.

Second, to complement the qualitative analysis outlined above with a quantitative assessment, we computed the key values of the identified topics. Keyness is a statistical indicator that measures the value, also known as the log-likelihood score (Rayson & Garside, 2000). This metric provides statistical meaning and makes it possible to measure the

Table 3
Relevant papers found in the literature review.

Authors	Purpose	Main topics	Main concepts analyzed and linked to the present study
Al-Mushayt (2019)	To propose an AI techniques framework and review models with the most recent advances for the improvement of e-government services	Artificial Intelligence, deep learning, sentiment analysis, smart e-government platform, e-government information management framework, trust, transparency, efficiency, IoT, Big data, *E*-Government Development Index (EGDI)	■ Application of AI in government platforms ■ Improving results in governments' objectives with the use of AI ■ Optimization of e-government services using AI
Androutsopoulou, Karacapilidis, Loukis, and Charalabidis (2019)	To propose a novel approach based on advanced chatbots to address communication and interaction challenges between citizens and government in a complex, ambiguous, and uncertain context	Media Richness Theory (MRT), ICT infrastructure, human and machine intelligence, digital channels, chatbots, natural language processing, machine learning, data mining technologies, data management services, knowledge processing services, application services	■ Understanding new technologies that work with AI in governments and their linkage to BDS ■ Exploring the interaction between governments and citizen in different environments
Ashok, Madan, Joha, and Sivarajah (2022)	To develop an ethical framework for AI and Digital technologies	Digital ethics, impact of AI on society, socio-technical transformation, AI principles, ethics in the application of AI	■ Identifying ethical considerations in relation to the adoption of AI ■ Classifying the main domains of ethical research in relation to AI and its uses ■ Identifying the main factors of influence of AI in relation to ethical, legal, social, and economic indicators ■ Exploring indications for the development of intelligent systems to gain trust and better adoption of AI by citizens
Benefo et al. (2022)	To analyze ethical, legal, social, and economic implications of the global application of artificial intelligence	Globalization, political factors, sociocultural analysis, ethical application of AI	
Biros (2020)	To examine the main challenges and vulnerabilities that new technologies present related to security, privacy and ethics	Artificial intelligence, Big Data analytics, Internet of Things, information security education, ethics, governance, and privacy	■ Identifying privacy vulnerabilities for users in AI ■ Analyzing the link between governance and privacy and the use of techniques and tools that work with AI.
Chamola, Hassija, Gupta, and Guizani (2020)	To explore how advanced technologies might help to deal in times of crises or pandemics by the improvement of processes and management in the public sector	Epidemics, Covid-19, global economy, Internet Medical of Things (IMoT), Internet of Things, healthcare IT systems, telemedicine, drone technology, wearables, GPS, GIS, Bluetooth, Artificial Intelligence, machine learning, Blockchain, 5G network technology	■ Classifying technologies that can be used for BDS by governments ■ Mapping government decision making during crises and pandemics and identify whether AI is used or not ■ Understanding the functioning of public sector decision making and the existence of BDS actions
Chatterjee (2019)	To study citizens' willingness to use robots when a strict artificial intelligence regulatory control is enforced	AI policy, ethics, regulations, artificial intelligence regulation, perceived ethical dilemma, perceived risk, control beliefs, quality of life of citizens, impact of AI regulation	■ Exploring AI regulatory policies in government ■ Identifying dilemmas linked to ethical and privacy issues ■ Assessing citizen's opinions about technologies that work with AI and their possible linkages to BDS
Chatterjee (2020)	To review how AI policy might be framed and developed by governments	AI opportunity, AI policy framework, government actions, adoption of AI, AI and challenges, AI policy recommendations, normative and responsible AI development, research and applications, acceleration adoption of AI, training and skilling	■ Understanding the future of AI regulation by governments ■ Identifying recommendations and challenges of AI and its adoption by governments and citizens in relation to BDS
Chatterjee and Sreenivasulu (2019)	To analyze how AI technologies and government regulations influence on sharing personal data and deal with human rights violations.	Personal data sharing (PDS), Human Right Abuses (HRA), Influence of AI for Analyzing and Profiling (IAAP), Influence of Regulation and Governance (IRG)	■ Exploring possible human rights violations linked to the use of BDS and AI techniques by governments ■ Analyzing the privacy of users' personal data
Chatterjee, Khorana, and Kizgin (2021)	To analyze citizen satisfaction in relation to the use of AI in governments	AI services, AI adoption, citizen satisfaction, operational and strategic public value for citizens	■ Classifying the services provided by governments when AI is used ■ Identification of citizen satisfaction for the theoretical assimilation of public value when AI is used ■ Developing uses of AI by governments to generate public services
Chen and Wen (2020)	To study people's attitude and trust towards governments and corporations depending on their perceptions of AI.	AI, Institutional trust, Human-Machine Communication (HMC), science trust, media, and news	■ Measuring the trust and attitudes towards AI regulatory policies that may affect BDS ■ Understanding the perception of AI techniques from the point of view of citizens and their use by governments ■ Exploring the ability of corporations to influence BDS through surveillance capitalism
Di Vaio, Hassan, and Alavoine (2022)	To develop a bibliometric study to understand the effectiveness relationships between data intelligence and human-artificial intelligence	Data intelligence, data analytics, public sector-decision making, decision-making effectiveness	■ Investigating how data intelligence improves decision making in the public sector

(*continued on next page*)

Table 3 (*continued*)

Authors	Purpose	Main topics	Main concepts analyzed and linked to the present study
Engin and Treleaven (2019)	To overview the current global applicability of data science automation by governments and to suggest new systems for the improvement of public services for citizens	Automation of government services; government data facilities, IoT, AI, Big Data, Behavioral/predictive analysis, Blockchain technologies, GovTech, public services, supporting civil servants, public records, national physical infrastructure, laws, statutes and compliance, public policy development, challenges and public debate.	■ Identifying the main emerging technologies related to improving performance in decision making ■ Analyzing whether data automation and government predictions and linked to BDS and surveillance capitalism ■ Identifying new ways to improve services for citizens and their security and privacy ■ Understanding the ways of predictive analysis in the Big Data era from governments perspectives
Furumura et al. (2020)	To explore large databases of past catastrophic earthquakes by using AI image recognition technology to predict potential disasters and their consequences	Data retrieval system, HERP database, digitalized images, scan seismograms	■ Analyzing a case of AI application on historical event databases ■ Understanding how the use of AI models and data prediction can be used in different sectors for BDS actions
Gonzalez, Ferro, and Liberona (2020)	To study the global use of AI in smart cities for the applicability of its advantages and benefits in other cities and to propose a model for its adoption	Smart city, smart cities models, intelligent economy, intelligent environment, intelligent government, intelligent life, intelligent mobility, intelligent people, AI applied to public transport, electricity supply, waste and paper management, health, security, digitalization	■ Understanding the case of AI in smart cities for data collection by governments ■ Measuring whether the implementation of connected cities can influence the actions linked to surveillance capitalism and prediction of citizens' behavior and their decisions
Haug et al. (2020)	To analyze the effectiveness of Non-Pharmaceutical Interventions (NPI) to alleviate the spread of Covid-19 by proposing a model combining computational techniques.	Statistics, inference, AI, nodes, data, regression techniques, ranking of NPI	■ Analyzing new computation techniques for the prediction and measurement of efficiency in different sectors ■ Understanding the use of AI in pandemics or disasters and its future use by governments and corporations
Hollander and Icerman (1991)	To evaluate how new technologies and expert systems can be effectively used in governmental financial planning and decision-making processes	AI technology, expert systems, feasibility analysis, capital budgeting decision process, political and economic considerations, cost-benefit analysis	■ Analyzing political and economic decision of governments with the use of new technologies and efficiency systems based on AI ■ Identifying decision-making processes in governments when using AI ■ Exploring economic decisions and its relationships with BDS and surveillance capitalism actions
Jimenez-Gomez, Cano-Carrillo, and Falcone Lanas (2020)	To highlight key aspects and foundations triggered by AI application by public institutions	Data-driven digital government, public value, AI-based systems, e-government, smart governance, data science, ethical and societal implications, digital technologies, interoperability of digital government, machine learning, deep learning, human rights, fundamental principles of good government, cybersecurity	■ Classifying government actions when using AI in decision-making ■ Understanding social and human rights implications from the governments' perspective
Kankanhalli et al. (2019)	To propose a research agenda on IoT and AI for the smart governments based on available literature and to identify the main challenges of these technologies.	IoT, AI, smart government, IoT-enabled AI systems, domains of smart governments, regulation and policy, AI and IoT principles, stakeholders	■ Identifying the challenges in relation to regulation and policy of IoT technology in governments ■ Measuring the degree of risk and privacy of citizens due to eavesdropping policies with IoT devices
Martín and León (2015)	To find new ways of interaction between citizens and e-governments by proposing a comprehensive approach that identifies relevant information through expert system technologies	Semantic web, AI, expert system technologies, semantic metadata, e-government	■ Understanding new options for interaction between citizens and governments with the use of systems that work with AI ■ Measuring the degree of acceptance of e-governments and their risks for citizens
Nasseef, Baabdullah, Alalwan, Lal, and Dwivedi (2021)	To analyze public healthcare management system with the use of AI to improve decision making	AI in governments, healthcare, improve decision-making, data-centric processes, knowledge, AI	■ Examining the effects of AI strategies used by governments in the healthcare sector ■ Developing a cognitive model for the theoretical study of AI in governments ■ Identifying and developing knowledge-based exchange practices
Pencheva, Esteve, and Mikhaylov (2018)	To review the role that Big Data and advanced analytics play in the public policy and administration field	Big Data, advanced analytics, public administration, policy cycle, agenda-setting, policy formulation, policy implementation, policy research, policy evaluation, system-level barriers, organizational-level barriers, individual-level barriers	■ Analyzing actions linked to Big Data in governments that may cause privacy violations for users ■ Understanding how predictive analytics can be used in the public sector
Polat and Alkan (2020)	To study the Good Governance approach addressed in the context of Government 3.0 in the land administration field.	Good Governance (GG), General Directorate of Land Registry and Cadastre (GDLRC), Land Administration System (LAS), property rights, land registry web application, Spatial Real Estate System (SRES), Land Registry Archive	■ Exploring uses and application of new technologies in smart governance ■ Analysis of the different use cases of information and prediction systems in public administration

(*continued on next page*)

Table 3 (*continued*)

Authors	Purpose	Main topics	Main concepts analyzed and linked to the present study
		Information System (LRAIS), Map Data Base (MDB)	
Shneiderman (2020)	To discuss the ethical principles of human-centered AI (HCAI) by an effective governance practice	Human rights, corporate social responsibility, the Governance structure for Human-centered AI, software engineering practices, reliable systems, safety culture, trustworthy certification	■ Ethical use of AI tools that can be used for BDS and surveillance capitalism by governments ■ Understanding the responsibilities of governments for the use of AI in relation to privacy and citizens' rights
Silva, Jian, and Chen (2015)	To propose a hybrid process analytics approach to achieve high operational efficiencies and high-quality assignments splitting up complex data into more manageable for R&D project selection	Hybrid process analytics approach, data-driven process models, clusters proposal, high-quality assignment, high-quality reviews, sematic language models, social learning theories, social network analysis, bibliometric analysis	■ Understanding new processes focused on data-driven models and their efficiency ■ Understanding how hybrid data models applied to social networks can be vulnerable in terms of information security breaches
e Silva., Rodrigues, and Ishii. (2020)	To propose a model based on AI technologies to reduce Infant Mortality Rates (IMR) in health systems	AI, Knowledge Discovered Data (KDD), Mortality Information Systems (MIS), Live Birth Information System (LBIS), RIGOR approach.	■ Analyzing new AI use to predict and optimize results in the healthcare industry ■ Understand whether new AI application and models could be used for behavior modification of citizens
Skaug Sætra (2020)	To examine the dangerous consequences of implementing AI by governments	AI technocracy, expert rule, democracy, political decision-making, AI, algorithm governance, transparency, Explainable AI (XAI), the magical decision box, AI supremacy, legitimacy, people participation, machines, morality and human well-being, transparency, accountability	■ Identifying and classify techniques and uses of AI by governments and their possible consequences for citizen privacy ■ Linking the consequences of the use of algorithms in governments to behavior modification and actions of both BDS and surveillance capitalism
Stoica, Pitic, and Mihăescu (2013)	To propose a novel model to analyze social media data for the development of e-business and e-governments	e-business, e-governments, social media, sentiment analysis, tweets collection, training data, topic classification	■ Classifying models that work with AI for both corporations and governments, and their applications in social networks ■ Understand the information flow of AI models in e-business and e-governments as well as their risks
Susanto, Yie, Rosiyadi, Basuki, and Setiana (2021)	To study the management of automations and AI in governments	Data security, connected governments, automation management, AI management in public institutions	■ Identifying management techniques used by governments to improve decision making with automation and AI ■ Analyzing the main risks related to data security in an era of connected governments
Wilson (2022)	To analyze national strategies related to AI and public engagement	Engagement values, technology frames, AI complexity, AI regulation, policymaking and AI applications	■ Analyzing public strategies in relation to the services offered to citizens ■ Identifying the main values to increase engagement in national strategies that use AI
Wong (2019)	To analyze data, algorithms, and machine-learning techniques for the development of smart cities. Through a speculative case study, to analyze possible challenges to deal with	Smart cities, Big data, algorithms, machine learning, data determinacy and fallacy, collective civic intelligence, surveillance, predictive behavior	■ Understand the consequences of the use of predictive algorithms and data collection in smart cities ■ Linking data intelligence with predictive actions and behavior modification ■ Identifying surveillance capitalism and prediction behavior and their possible use by corporations and governments
Zato, De Luis, Bajo, De Paz, and Corchado (2011)	To propose a hybrid AI system to reduce resources and increase profitability	Hybrid AI system, multi-agent systems, hard type, virtual organizations (VO), case-based reasoning systems (CBR), planning tasks, task assignment	■ Understanding how AI models can be linked to economies of scale actions in which profitability is the major indicator ■ Linking these models to strategies in order to increase profitability in governments
Zheng et al. (2020)	To propose an automated and agile platform to enhance efficiency and reduce moral hazard in civil servants' work environments.	Civil servants, moral hazard, policy, government service provision, automatization	■ Exploring the use of models based on data-centric systems and platforms to predict behavior ■ Understanding the use of AI systems in different industries and their benefits for governments

Source: The authors.

relevance of different topics in the same database or corpus. According to Duran, Hall, McCarthy, & McNamara (2010), the log-likelihood score of 3.8 or higher was reported to be statistically significant at $p < 0.05$. Therefore, the interview conversations were established as inputs phrases, and text documents were considered as sub-corpora of the original corpus. Statistical significance in this study was considered when $p < 0.05$ Drmota, Szpankowski, & Viswanathan (2012).

Furthermore, we used textual analysis computed in Python (Anand, Bochkay, & Chychyla, 2020). With this approach, it is possible to identify values in the form of insights using in-depth content analysis (Millstein, 2020). Specifically, the variables related to the weighted percentages/frequency of a keyword in the database composed of the set of interviews were studied (McHugh et al., 2020). In this way, the relevance of certain keywords was obtained (Auer, 2018). Based on the percentages of relevance achieved, we established parameters that casually explained the objectives of the present study (Saura, Ribeiro-Soriano, & Palacios-Marqués, 2021b). This exploratory approach follows the indications of content analysis using the NLP framework.

An analysis of the main n-grams collected in the coded text of the interviews was also performed. In order to compute the n-grams

Table 4
Interviewees by role of informant, industry, professional state, organization, and nationality.

In.	Informant's position	Industry	Professional Status	Organization	Nationality
A	Government economist in Spain	Economy	Retired	Government	Spanish
B	Manager at the European Investment Bank	Politics, Economy	Active	Government	Spanish
C	Ambassador of Spain I	Politics, Diplomacy	Retired	Government	Spanish
D	Ambassador of Spain II	Politics, Diplomacy	Retired	Government	Spanish
E	Ambassador of Spain III	Politics, Diplomacy	Retired	Government	Spanish
F	Ambassador of Venezuela	Politics, Diplomacy	Active	Government	Venezuelan
G	Business Confederation of Madrid (CEIM) *	Economy	Active	Private Organization	Spanish
H	PSOE* Senator	Politics	Active	Government	Spanish
I	Regional Representative PP*	Politics	Active	Political party	Spanish
J	Senator of the PP I	Politics	Active	Political party	Spanish
K	Senator of the PP II	Politics	Active	Political party	Spanish
L	Mayor of a city in Galicia	Politics	Active	Political party	Spanish
M	PP Town Councilor	Politics	Active	Government	Spanish
N	Member of League of Arab States*	Politics, Economy	Retired	Private Organization	Egyptian
O	Consul of Colombia	Politics, Diplomacy	Active	Government	Colombian

*Business Confederation of Madrid (CEIM) advice the national government of Spain.
*League of Arab States is a regional organization that aims to safeguard independence and sovereignty and to consider in a general way the affairs and interests of the Arab countries.
*PP (Partido Popular) is a Spanish political party. Currently, it is the opposition party that controls 6 of 19 regional governments in Spain.
*PSOE (Partido Socialista Obrero Español) is the Spanish political party in the current government.

Table 5
Demographic characteristics of the informants.

Demographic characteristic	Specification	Count	(%)
Gender	Female	4	26,6
	Male	11	73,3
Profession	Politician	6	40
	Government economist	2	13,3
	Diplomat	5	33,3
	Government advisor	2	13,3
Education	Postgraduate	12	80
	PhD	3	20
Age	36–45	4	26,6
	46–56	10	66,6
	57–67	1	6,6
	>68	1	6,6

analysis, we followed Wu and Su (1993) who argued that statistical analysis of the measure known as mutual information (MI) is justified when using textual analysis and n-grams. This indicator refers to the probability of co-occurrence of two variables that are correlated. Likewise, Bouma (2009) and Iyengar et al. (2012) used MI indicator between random variables X and Y. Of these, those with marginal probabilities and p(x) and p (y), and joint probabilities p (x, y), can be computed.

4. Results

4.1. Results of systematic literature review

According to the results of the systematic literature review, in the studies included in the dataset, we identified the main uses that governments make of AI. In this way, the interviews were developed based on the results of this methodological process. Furthermore, to complement the results obtained through our systematic literature review as indicated previously, we conducted interviews with informants who work or have worked in governments. The interviews were based on the main concepts related to user privacy and governments' use of AI found in the literature (see Table 6). Therefore, the aim of the interviews was not only to understand new uses of AI, but also to obtain information regarding user privacy and how user information is treated based on the results presented in Table 6 from the literature review.

Regarding the major identified uses of AI by the governments, the main one is the continuous development of new models that increase the efficiency of the results (Chamola et al., 2020). This is a characteristic of AI, since the more the models that work with machine learning are trained, the greater the efficiency in terms of prediction of finance is, if the objects are focused on profitability. Likewise, the uses focused on decision making for process improvement and the evolution of management and governance practices were also remarkable (Skaug Sætra, 2020).

In this way, techniques are used to understand and optimize interactions with citizens (Androutsopoulou et al. (2019) through channels such as social networks (Saura, Palacios-Marqués, & Iturricha-Fernández, 2021; Silva et al., 2015), as well as information systems or

Table 6
Main uses of AI by governments found in the literature review.

N.	Artificial Intelligence deployment	Authors
1	AI application to the improvement of e-government services	Jimenez-Gomez et al. (2020); Martín and León (2015); Stoica et al. (2013)
2	The use of AI on smart e-government platforms	Zheng et al. (2020); Al-Mushayt (2019)
3	Chatbots development with AI to address communication and interaction challenges between citizens and government	Androutsopoulou et al. (2019)
4	Use of AI for the development of knowledge processing services	Chamola et al. (2020)
5	AI application to improve government research	Kankanhalli et al. (2019); Pencheva et al. (2018); Chatterjee (2020)
6	AI to improve the training and skilling of governments' platforms	Chatterjee (2020); Chamola et al. (2020)
7	Use of data science automation by governments on their platforms	Chen and Wen (2020); Engin and Treleaven (2019)
8	Use of AI for the improvement of public services for citizens	Gonzalez et al. (2020); Pencheva et al. (2018)
9	Prediction of events and happenings based on the database analysis of other events using AI	Polat and Alkan (2020); Wong (2019)
10	AI to help governmental financial planning and decision-making processes	Zato et al. (2011); Skaug Sætra (2020)
11	Use of AI in IoT to boost the development of smart governments	Al-Mushayt (2019)
12	Development of a novel model to analyze social media data for the development of e-governments	Stoica, Pitic, and Mihăescu (2013); Silva, Jian, and Chen (2015)
13	Development of a hybrid AI system for the reduction of resources and increase of profitability in government systems	Zato, De Luis, Bajo, De Paz, and Corchado (2011); Silva, Jian, and Chen (2015)

Source: The authors.

data exchange platforms. Automation and the use of models and algorithms are being increasingly widespread in governments, as they, through new technologies linked to SI (chatbots, IoT, smart cities, among others), try to collect databases that can predict how society is organized, determine financial models, and improve the optimization of industries and cities (Silva et al., 2015; Zato et al., 2011).

Similarly, in order to cover the objectives proposed in the present study, Table 7 details the main privacy issues for users and citizens, and concepts linked to AI uses found in the literature review. Of note, concepts linked to the use of AI and security of user data, prediction, and analysis of their behavior, as well as its modification, were considered (see Zuboff, 2019a, 2019b).

The results of our analysis of privacy issues and concepts found in the literature review and presented in Table 7 highlights the ease with which governments have access to citizen data to train AI models (Engin & Treleaven, 2019; Wong, 2019). Precisely, public institutions try to solve this fact with the initiatives for good governance (Martín & León, 2015). However, citizen privacy is a human right directly linked to the legitimate use of data and access to citizen information (Chatterjee & Sreenivasulu, 2019). Predicting citizens' behavior based on the data they generate, in economic, social or health terms, is relatively easy with the numerous data analysis techniques that use AI to make predictions (Biros, 2020).

The problem lies mainly in the data protection regulations that may allow government to use these techniques without violating citizens' privacy, as highlighted by Shneiderman (2020). In this way, a balance must be found between the use of citizens' data to make predictions of their behavior by governments. This can be done by improving economic, social, or cultural indicators (Polat & Alkan, 2020; Skaug Sætra, 2020). If governments develop strategies focused on profitability indicators, citizens' data become economies of scale that can lead to illicit BDS practices to modify, whether intentionally or unintentionally, citizens' behavior (Ribeiro-Navarrete et al., 2021).

4.2. LDA and textual analysis of interview data

Using the LDA process, a total of 7 topics were identified. Of these, 4 topics were related to privacy issues (Human behavior, Behavioral predictions, Data privacy law and regulation, and Risk of behavior modification) and 3 further topics were related to AI deployment by governments (Intelligence decision making, Digital surveillance and Decision automation). Table 8 summarizes the identified topics, their descriptions, and the corresponding indicators of keyness and *p*-value.

Based on the results of textual analysis, the most frequent words are presented in Table 9. In addition to measuring the weighted percentage in the entire database, the keywords were grouped by similarity.

To obtain additional insights using data mining, we also defined n-grams supported in placement analysis that takes into account the contexts where words occur in a corpus (Biber, 2004; McEnery & Hardie, 2013). In this way, we analyzed the position of the main words in the database, with a particular focus on the place where a word is positioned. Therefore, placement presents a strong and stable relationship, also called a lexical or n-gram package. Table 10 lists the identified n-grams presented by rank (R), with the words identified in Table 9.

Here, frequency refers to the total frequency of appearance of the collocates in the in-depth interviews database. As indicated in Saura, Ribeiro-Soriano, & Iturricha-Fernández (2022), this is the sum of Freq L of the words that appear on the left on the topic and Freq R of the words that appear on the right of the topic.

5. Discussion

In the present study, we explored the main uses and techniques of AI developed by governments, as well as investigated the main concerns related to user privacy. Our analysis of the qualitative interviews analyzed using several data-mining techniques yielded several important insights.

Overall, governments consider knowledge of citizens' behavior to be key for the success of good governance (Chatterjee, 2019). However, as demonstrated in previous research on human behavior applying AI techniques, both predictions and correlations that can be identified in the collective behavior analysis pose serious risks to user privacy (e.g., Biros, 2020).

Furthermore, the literature review process provided a thorough understanding of the main research developed in these fields, thus getting 13 uses related to AI in governments and 11 issues related to citizens' privacy. As stated by Zuiderwijk et al. (2021), these insights can be used to outline the interviews as an additional method, as well as to create theory and knowledge in relation to the studied topic.

Similarly, in relation to the identified interview topics related to the predictions of behavior, it becomes clear that both the feelings and the

Table 7
Main privacy issues and concepts found in the literature review.

N.	Artificial intelligence and privacy issues	Authors
1	Analysis of human behavior and study of machine intelligence linked to privacy issues	Wong (2019); Engin and Treleaven (2019)
2	Information security education linked to ethics, governance, and privacy	Biros (2020); Shneiderman (2020)
3	The role of user data and information in the global economy	Chamola et al. (2020)
4	Development of AI regulations and ethics	Chatterjee and Sreenivasulu (2019)
5	Study of an AI policy framework for government actions	Kankanhalli et al. (2019); Chen and Wen (2020)
6	Study of normative and responsible AI development by government.	Polat and Alkan (2020); Skaug Sætra (2020)
7	Privacy issues in personal data sharing and human rights abuse	Chatterjee and Sreenivasulu (2019); Chatterjee (2019)
8	Good governance when using AI and user data.	Martín and León (2015); Polat and Alkan (2020);
9	Structure of governance when developing human-centered AI analysis	Pencheva et al. (2018); Shneiderman (2020);
10	Software engineering practices in governments and its influence on data privacy	Gonzalez et al. (2020); Polat and Alkan (2020);
11	Development of reliable AI systems, safety culture, trustworthy certification to preserve society privacy	Kankanhalli et al. (2019); Biros (2020)

Source: The authors.

Table 8
Topics identified using LDA.

R	Topics	Topic description	Keyness	*p*-value
1	Human behavior	Study and optimization of the understanding of society and its behavior	946.26	0.045
2	Behavioral predictions	Predict citizens' behavior, including their feelings, movements, actions, criminal acts, or cyber-attacks	900.05	0.042
3	Intelligence decision-making	Decision-making focused on data analyzed with AI in governments.	870.13	0.038
4	Decision automation	Automation of government decisions in critical or alarm situations	869.93	00.38
5	Digital surveillance	Digital surveillance of the population based on the analysis of their digital and Internet-generated data	714.91	0.026
6	Data privacy law and regulation	Development of privacy law rules and their regulation	714.84	0.026
7	Risk of behavior modification	Manipulation of the company based on the automated study of the data using AI	302.08	0.010

Source: The authors.

Table 9
Keywords grouped by relevance.

Keywords	Freq.	WP
Human behavior, behavioral patterns, behavioral predictions, users' behaviors online, etc.	124	4.89
Intelligence decision-making, decision automation, smart decisions, artificial intelligence decisions, etc.	83	3.29
Data privacy, privacy regulation, data privacy law, data privacy policy, data privacy abuse, etc.	79	3.03
Digital surveillance, digital surveillance and privacy, surveillance capitalism, surveillance systems, etc.	75	2.99
Risk of behavior modification, user behavior predictions, behavior modeling, online users' behavior, etc.	67	2.83
Law and regulation, data protection, general data regulation, data policy, data policy protection, etc.	60	2.74

Source: The authors.

Table 10
N-grams analysis per identified topic.

R	Collocates for "Human behavior"			
	Freq	Freq L	Freq R	Collocate
1	49	23	23	BehaviorAnalysis
2	32	13	19	BehavioralPredictions
3	29	11	17	UsersBehaviors
4	19	9	10	BehavioralAnalytics

R	Collocates for "Behavioral predictions"			
	Freq	Freq L	Freq R	Collocate
1	38	23	15	Optimize
2	32	18	14	Actions
3	30	9	21	Surveillance
4	19	7	12	Alarm

R	Collocates for "Intelligence decision making"			
	Freq	Freq L	Freq R	Collocate
1	29	14	15	Intelligence
2	17	9	8	Systems
3	14	4	10	SmartActions
4	10	7	3	Governance

R	Collocates for "Decision automation"			
	Freq	Freq L	Freq R	Collocate
1	17	9	8	Decisions
2	14	6	8	GoodGovernance
3	10	4	6	Analysis
4	8	2	6	Understand

R	Collocates for "Data privacy law and regulation"			
	Freq	Freq L	Freq R	Collocate
1	34	15	19	DataRegulation
2	31	13	18	Privacy
3	30	17	13	Law
4	11	6	5	Abused

R	Collocates for "Risk of digital manipulation"			
	Freq	Freq L	Freq R	Collocate
1	26	17	9	Surveillance
2	15	6	9	Manipulation
3	14	4	10	Modification
4	6	2	4	Systems

actions derived from the analyzed data can be studied to optimize processes or to control the population. This suggests that there is a risk that governments can develop actions linked to surveillance capitalism or perform illegitimate actions of collective behavior analysis (see Zuboff, 2019a, 2019b), if these are used without caution. However, as indicated by Kamolov and Teteryatnikov (2021), these actions are powerful tools that can drive smart and good governance.

Although behavioral reactions can be meaningfully used to prepare for possible states of alarm, there is also a need to explore the ways to prevent, for example, cyberattacks that may jeopardize the user privacy; similarly, there is an urgent need to explore the limits of privacy when studying how the population will act (Engin & Treleaven, 2019). In this relation, Informant I pointed out to the following issue: *"We use artificial intelligence to predict possible criminal acts in the city. When artificial intelligence and our analyses tell us that there is a neighborhood where serious crimes, such as murder, can be committed, we increase the number of police patrols in those neighborhoods and with this, we try to act more quickly"*.

Therefore, as mentioned by the aforementioned interview participant, although governments can effectively use AI techniques to prevent illegal actions (Jimenez-Gomez et al., 2020), when AI is embedded in government strategies and decisions are automated, such as in critical or Covid-19 pandemic alert situations (Chamola et al., 2020), the risk of abusing user privacy increases (Chatterjee & Sreenivasulu, 2019), although from the government's point of view, it optimizes decision-making processes and data-driven decisions (Ribeiro-Navarrete et al., 2021).

A similar point was made by Informant M: *"In states of alert such as that generated by the COVID-19 pandemic, the use of artificial intelligence to predict possible infections and deaths has been used with statistical models. These models have helped us to both improve health care and the movement of people in cities, when a lockdown has been necessary."* The participant added, *"but the use of applications to track the location of user devices, although always anonymously, has highlighted the need to regulate the use of both artificial intelligence technology and other similar technologies to control the population in some way*." These indications contrast the results reported by Zhu, Chen, Dong, and Wang (2021) in their study in China, where the control and prevention of pandemics or diseases with the use of AI becomes a priority to achieve governments' aims, while the alert state can justify the performed actions. From the quote above, we can conclude that from the government perspective, it is taken into account that privacy is a powerful strategy that can be used for digital surveillance.

In today's digital era, the data generated by citizens can help anticipate their movements also from the marketing perspective, as in inducing users buy products or services (Martín & León, 2015) through manipulation on the Internet; this is typically done through an analysis of people's in customer journey (Dwivedi et al., 2020), online decision-making, or creating addiction in unethical strategies in social networks.

Furthermore, Informant B stated that *"Intelligence in governments has been used for several years. We focus mainly on listening and predicting possible causes affecting the State. However, it is true that there is still a wide range for the development of privacy and regulatory standards, and how these technologies and their applications can try to use user data, respecting or not their privacy"*.

However, one of the challenges in terms of surveillance of the society is to understand how governments can implement AI from the point of view of automatic decision-making (Chatterjee & Sreenivasulu, 2019). The prediction of user behavior is determined by the source of the data, which, in turn, can lead to digital manipulation of users (Stoica et al., 2013). Citizens should be aware of how governments will use their data and authorize (or not) the use of their data to train predictive models to, for instance, anticipate their movements and locations (Ribeiro-Navarrete et al., 2021).

Specifically, if AI is used by governments with a focus on making smart decisions, as is the case of the aforementioned informant, the risk to the privacy of users' information is lower (Jimenez-Gomez et al., 2020). However, when AI is consolidated in government strategies, and when decisions are automated, such as in critical situations or in a state of alarm due to the Covid-19 pandemic (Chamola et al., 2020), the risks

to user privacy violations increases (Chatterjee & Sreenivasulu, 2019).

Therefore, if governments have access to third party data, and these are linked to their national intelligence which already has access to massive data, this could led to possible manipulation and decision making focused on surveillance capitalism (see also Shneiderman, 2020).

In this respect, Informant N made the following observation: *"Governments have access to a multitude of sources of data on citizens and users. However, governments always use legitimate sources of information and a priori, they do not have access to third-party sources that can pass on personal data of users to governments for use in non-legitimate artificial intelligence models."* He then continued: *"And concerning the risk for manipulation of citizens and surveillance, national security processes increasingly use artificial intelligence, and as we know it works with data. The risk of manipulation does not exist because citizens are free in their actions and artificial intelligence and automation is intended to predict how the tasks that the government performs can be optimized and are always legitimate."*

In this way, Informant N highlights the use of AI tools and strategies in the government for the management of decision making, optimization of messages and conversations with citizens, as well as automation in public domain decisions (see also Chen & Wen, 2021). As argued by Zuboff (2019b), the power of prediction through access to millions of data can cause systematic violations of citizens' privacy, even without governments being aware of it due to a misunderstanding of the technology (see also Saura, Ribeiro-Soriano, & Palacios-Marqués, 2021a). Therefore, the automation of decisions focused on AI by governments should be regulated (Engin & Treleaven, 2019). Although governments attempt to use legitimate data sources, there are already some examples when, despite the legitimate intentions of the government, the companies that passed those data to governments had made illegitimate use of those data (Thompson & Warzel, 2019).

Accordingly, the information generated by users both on the Internet and on digital devices must comply with a new regulatory framework for data protection. Users must have the right for their data and decide whether or not these data could be transferred to companies (Pencheva et al., 2018). However, at present, the data are a currency so that if, for instance, users want to use an application, they a priori accept the privacy policy and, in case it is rejected, they will not be able to use that application. As indicated by Caudill and Murphy (2000), this can be understood as blackmail of users; indeed, in most cases, the option of buying an application with the option of not giving the data to the company, which may subsequently sell those data, does not exist (Bennett & Raab, 2020). Yet this and other initiatives (Obar & Oeldorf-Hirsch, 2020) may allow one to predict user behavior and use BDS without undermining collective behavior analysis.

In addition, with respect to surveillance capitalism and the use of AI by governments, specific regulation should ensure that the used data sources are legitimate and that they are not used, consciously or unconsciously, to manipulate the population so that to obtain economic benefits from both the government and the companies working with the data.

5.1. Future research agenda

The development of governmental uses and practices of AI has been defined to be essential for the future of governance that supports new technologies (Chatterjee et al., 2021). The adoption and use of these technologies should focus on improving services to citizens and society in general. However, the challenges and risks of new forms of AI need to be properly understood and studied in the future (Chen & Wen, 2021). Accordingly, the identification of different techniques developed by governments for the acquisition and collection of massive data from citizens becomes a priority. As argued by Zuboof (2019a), the ethical principles and values that ensure the privacy of citizens should be properly defined and classified so that governments can establish good practices in the future.

Applying AI in the processes developed by governments can help to predict the behavior of citizens. Accordingly, regulations must be developed so that governments can make legal use of tools to predict society behavior (Wilson, 2022). Decision-centric tools working with AI and data automation must draw the line between decisions that need to be made by humans and machines. As indicated by Al-Mushayt (2019), the automation of data analysis and behavior prediction algorithms must comply with regulations that ensure legitimacy of user privacy (Ashok et al., 2022). Parallel to these new processes, new limits must be established to avoid the risk that, through initiatives to modify mass behavior, governments can achieve non-legitimate or non-lawful objectives promoted by states of emergency or national security nature (Susanto et al., 2021). Therefore, the development of the influence of surveillance capitalism on the uses and practices of DBS techniques should be studied in depth (Zuboff, 2019b).

In the present study, we observed that the existing relationships between user privacy, risk of personal data management, and promotion of actions that can modify citizens behavior are just some of the challenges that researchers should study in depth in the medium and long term (Benefo et al., 2022). In addition to automation and the exponential development of AI, these new technologies must be regulated in advance, as, while technology advances exponentially, legislation and its development entail longer time horizons (Di Vaio et al., 2022). Governments must be aware of this weakness and should understand and develop legislation related to AI and its possible unethical uses well in advance (Ribeiro-Navarrete et al., 2021).

Likewise, the efficiency in the performance of the AI-based strategies, as well as its risks and economic benefits, must be correctly defined and classified. If we closely attend to these issues, information and data processing and improvement of decision-making by governments (Nasseef et al., 2021) must be correctly designed in practice. From the perspective of citizens' benefits, and from their relationship with public institutions, these uses should be correctly implemented in government's strategies. Governments must ensure that society can trust the uses of AI in relation to massive analysis of behavior and collective intelligence.

Therefore, considering the points outlined above, seeking to structure the future of BDS exploration and to contribute towards creating collective behavior analysis strategies of society, we present an agenda of future research questions that must be answered in further research on AI implemented by governments and user privacy (see Table 11).

5.2. Theoretical implications

The first contribution of the present study is that our results bridge a gap in the literature that, until now, has lacked a thorough application of the concept of BDS to study decision making by governments using AI. Accordingly, our results can be used by other academics to design new research on user privacy, predicting user behavior, or optimizing decision making in governments. More specifically, this study provides theoretical information related to collective governance and the improvement of government services with the use of AI. In addition, our results suggest that governments must participate in the development and application of AI-related regulations. It must be understood and theorized that, through the development of public policy, easily measurable strategies and processes should be established. Decision-making deployment should focus on a safe improvement of AI applications. The development of AI must be linked to the legislation and regulation of government relations with third parties, as well as with the companies that collect the data and transfer those data to public institutions.

Similarly, ethical governance is a key element for governments to follow ethical practices and monitor the establishment of new AI functions to predict the behavior of society. Citizens' surveillance and the legitimate use of AI by governments close the cycle of identification and classification of the main uses and practices that governments perform

Table 11
Future research questions on user privacy and AI strategies deployed by governments.

Area of research	Key elements	Future research questions
Citizen's behavioral data collection	Massive data acquisition and collection	■ Is it ethical to collect and analyze non-intentionally generated data of citizens? ■ Will such analysis violate citizens' privacy? ■ How can users decide what data to pass on or not to governments for analysis?
Government AI uses	Application of AI in the activities and processes developed by governments	■ What are the limits of predicting citizen and user behavior when using AI? ■ What regulations should be in place to explain to citizens the outputs that can be obtained by their data analysis? ■ How can governments better understand machine behavior and algorithmic behavior?
Automatic decision making in governments	Decision making with support from AI dashboards	■ What is the regulatory framework for governments to make automatic data-based decisions using AI? ■ What should be the limits of the automatic analysis of citizens' data? ■ How can AI violate user privacy rights when predicting their behavior and movements?
Illegitimate uses of AI	Regulatory framework to ensure data legitimacy and user privacy	■ How should governments inform the population of the sources of user data acquisition and the corresponding uses? ■ How can AI violate user privacy rights when predicting their behavior and movements?
Citizens' behavior modification and prediction	Setting limits to the prediction of user behavior to avoid mass behavior modification	■ In collective behavior analysis, is it possible to individualize actions and predictions about a previously anonymous individual? ■ Can governments induce behavior modification without being aware that AI strategies can modify citizens' behaviors?
Citizen's digital surveillance	Optimization of surveillance processes that do not violate user privacy	■ How can governments ensure that users' digital surveillance is legitimate? ■ What is the regulatory framework to ensure that governments do not have access to personal data when using Internet data collection actions with AI? ■ How can we ensure that governments do not promote actions linked to the concept of surveillance capitalism?

to date in relation to AI.

In addition, from the perspective of collective behavior analysis, the present study theoretically discusses the contributions identified as research topics, which can be established both as constructs of quantitative models and as research objectives in further exploratory or qualitative research. In this way, the study contributions can be used to establish and design new approaches that use exploratory methods that work with AI to make predictions for this field of research. The emerging development of AI and the analysis of collective behavior linked to the privacy of citizens become relevant issues for the next decade, since AI applications and development will be exponential in this time horizon.

In addition, our findings from the interviews with members of the government allowed us to identify the main concerns related to user privacy and use of AI from the point of view of both governments and citizens. The present study also contributes to the ongoing debates about surveillance capitalism, and how user data can become the core of economic initiatives and stimulations of the global economy. Furthermore, through the development of interviews, we interpreted the informants' points of view about AI and its uses in public institutions, as well as linked the results to the main initiatives and applications of AI developed to date, in relation to both economic impulses and predictions of mass behavior.

5.3. Implications for governments

The results of the present study provide several practical implications for government. First, governments can use our findings as a reference to the main uses of AI previously discussed in the literature. Also, based on our analysis of privacy and ethical issues, governments can take into account the future research agenda proposed in the present study in order to avoid possible data breaches, violations of citizens' privacy, or abuses in the handling of their data. Therefore, governments need to be aware of the challenges and risks of using AI and BDS techniques to predict societal behavior.

Furthermore, governments can use the results of the present study to better understand the main applications of AI and how they should both respect the source of data collection and ensure appropriate use of predictive tools that do not violate user privacy. In particular, governments can use the proposed research agenda as a roadmap for the development of AI strategies in their policies and interactions with citizens. The future questions presented in the research agenda should be considered by governments and public agents to regulate the AI industry, avoid the use of unethical actions linked to BDS, and, above all, better understand the concept of surveillance capitalism and how AI actions can violate citizens' human rights.

Finally, governments should deploy regulatory decisions regarding user privacy on the Internet, management of data, and legitimacy of the study of the collective behavior. Likewise, governments can consult the studies reviewed in the present research in order to analyze detailed case studies that report AI preliminary results on behavioral prediction, crime anticipation, prediction of economic and social movements or health alerts.

5.4. Limitations

The limitations of the present study are related to the methodological approaches we used. Furthermore, since the object of our investigation is a fast-developing field, some of our conclusions might eventually become outdated. Other limitations include a relatively small number of interviewed informants, as well as the fact that the research was limited to only one country (Spain). Of note, the study includes articles only in English, thus other valid research in different languages may be left out in the review. Also, the data mining approach with LDA is exploratory. Another limitation is that the names of the topics were chosen based on the results of exploratory research. Moreover, algorithms working with machine learning can improve their efficiency through training. Of note,

in future studies, it would be necessary to address research questions outlined in the proposed agenda in order to complete and answer the questions identified as priorities in the given research topic.

6. Conclusions

In the present study, we explored the concept of BDS linked to privacy issues when governments develop strategies using AI. To this end, we performed a systematic review of the literature and collected major academic contributions published to date. In addition, we conducted 15 semi-structured interviews with experts working in public administrations and governments, and two data-mining approaches were used to analyze the collected interview data. Based on the results, we formulated a future agenda for further research in the studied area.

With regard to RQ1 (*What kind of citizens' privacy issues are expected when governments use behavioral-based AI in their strategies?*), we classified the risks to citizens' privacy according to the types of strategies focused on AI used by governments. These issues were analyzed and incorporated into the proposed future research agenda. Furthermore, concerning RQ2 (*What AI techniques can governments develop to predict the society's behavior?*), we used our systematic literature review to find out the main uses that governments make of AI. Based on these findings, we also discussed possible applications from the perspective of collective behavior analysis.

In addition, with a particular focus on user privacy, we also defined different perspectives of analysis of user behavior and privacy violations. To this end, we identified several major topics in our data. With regard to the last objective of the present study, we established future guidelines to address challenges to conduct further research on different areas and applications of AI by governments, secure preservation of data, and user privacy. Therefore, our results revealed the main uses of AI by governments and how they, through using AI in their models and algorithms that work with machine learning, focus on indicators to improve the interaction with citizens, organization in cities, services provided or the economy.

Similarly, we discussed the importance that behavior of citizens' knowledge for the success of good governance. The main issues related to the use of AI and the process of population control, and its massive monitoring were critically reviewed. In addition, we also discussed activities that governments perform in states of alarm in relation to the prevention of possible terrorist attacks or cyber-attacks where governments can use AI tools to defend the interests of the country. Additionally, we critically reviewed and discussed the kind of actions and decision-making processes carried out by governments in states of exception. Using these tools, governments can promote the development of AI-based actions—justified based on social impulses—but that could not respect the privacy of citizens. Therefore, favorable economic factors for governments, public institutions, or interested third parties should be promoted. Finally, the role of national intelligence for the analysis of collective behavior and the initiatives that governments can implement to ensure that their actions are legitimate and that the population supports and understands them correctly were also highlighted.

Finally, the development of regulations focused on the ethical design of user/citizen data collection and management strategies is not progressing at the same speed that technology. This elicits serious concerns about privacy of users and abuse of citizens' behavior with BDS techniques. Governments must implement new actions focused on regulating the security, ethics, and privacy of users' data. The risk of modifying citizens' behavior is real, so new legislation and rules must be implemented to regulate the use of citizens' data based on the optimization of models, development of intelligent systems, and actions to optimize industries or services.

CRediT authorship contribution statement

Jose Ramon Saura: Conceptualization, Methodology, Writing – review & editing. **Domingo Ribeiro-Soriano:** Writing – original draft, Writing – review & editing, Supervision. **Daniel Palacios-Marqués:** Investigation, Writing – review & editing, Supervision.

Acknowledgements

In gratitude to the Ministry of Science, Innovation and Universities and the European Regional Development. Fund: RTI2018-096295-B-C22

Appendix A

Table A.1
Interview questions.

Questions	Codification
What are the benefits of behavioral data sciences (BDS) analysis in governments artificial intelligence deployment?	QD1
Is collective behavior analysis important to governments? Why?	QD2
What is the use that governments make of citizens behavioral data?	QD3
Do you know the concept of Surveillance Capitalism? How do you think it influences the government surveillance actions?	QD4
Do governments consider users' and citizen's behavioral patterns analysis as possible privacy vulnerabilities?	QD5
What is the future of user privacy in terms of legitimacy of collection and use of their data?	QD6

References

Abou Elassad, Z. E., Mousannif, H., Al Moatassime, H., & Karkouch, A. (2020). The application of machine learning techniques for driving behavior analysis: A conceptual framework and a systematic literature review. *Engineering Applications of Artificial Intelligence, 87*, Article 103312. https://doi.org/10.1016/j.engappai.2019.103312

Acar, G., Englehardt, S., & Narayanan, A. (2020). No boundaries: Data exfiltration by third parties embedded on web pages. *Proceedings on Privacy Enhancing Technologies, 2020*(4), 220–238. https://doi.org/10.2478/popets-2020-0070

Agarwal, R., & Dhar, V. (2014). Big data, data science, and analytics: The opportunity and challenge for IS research. *Information Systems Research, 25*(3), 443–448. https://doi.org/10.1287/isre.2014.0546

Akter, S., Bandara, R., Hani, U., Wamba, S. F., Foropon, C., & Papadopoulos, T. (2019). Analytics-based decision-making for service systems: A qualitative study and agenda for future research. *International Journal of Information Management, 48*, 85–95. https://doi.org/10.1016/j.ijinfomgt.2019.01.020

Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: A systematic review and agenda for future research. *Electronic Markets, 26*(2), 173–194. https://doi.org/10.1017/S0963180114000589

Al-Mushayt, O. S. (2019). Automating E-government services with artificial intelligence. *IEEE Access, 7*, 146821–146829. https://doi.org/10.1109/access.2019.2946204

Altman, M., Wood, A., O'Brien, D. R., Vadhan, S., & Gasser, U. (2015). Towards a modern approach to privacy-aware government data releases. *Berkeley Technology Law Journal, 30*(3), 1967–2072.

Anand, V., Bochkay, K., Chychyla, R., & Leone, A. J. (2020). *Using Python for Text Analysis in Accounting Research. Forthcoming, Foundations and Trends in Millstein, F. (2020). Natural language processing with python: natural language processing using NLTK.* Frank Millstein.

Andrew, J., & Baker, M. (2019). The general data protection regulation in the age of surveillance capitalism. *Journal of Business Ethics, 1-14*. https://doi.org/10.1007/s10551-019-04239-z

Androutsopoulou, A., Karacapilidis, N., Loukis, E., & Charalabidis, Y. (2019). Transforming the communication between citizens and government through AI-guided chatbots. *Government Information Quarterly, 36*(2), 358–367. https://doi.org/10.1016/j.giq.2018.10.001

Ashok, M., Madan, R., Joha, A., & Sivarajah, U. (2022). Ethical framework for artificial intelligence and digital technologies. *International Journal of Information Management, 62*, Article 102433. https://doi.org/10.1016/j.ijinfomgt.2021.102433

Auer, E. M. L. (2018). *Detecting deceptive impression management behaviors in interviews using natural language processing*.

Bacq, S., Janssen, F., & Noël, C. (2019). What happens next? A qualitative study of founder succession in social enterprises. *Journal of Small Business Management, 57*(3), 820–844. https://doi.org/10.1111/jsbm.12326

Ballestar, M. T., Camiña, E., Díaz-Chao, Á., & Torrent-Sellens, J. (2021). Productivity and employment effects of digital complementarities. *Journal of Innovation and Knowledge, 6*(3), 177–190. https://doi.org/10.1016/j.jik.2020.10.006

Belhadi, A., Djenouri, Y., Srivastava, G., Djenouri, D., Lin, J. C. W., & Fortino, G. (2021). Deep learning for pedestrian collective behavior analysis in smart cities: A model of group trajectory outlier detection. *Information Fusion, 65*, 13–20. https://doi.org/10.1016/j.inffus.2020.08.003

Bem, D. J. (1995). Writing a review article for psychological bulletin. *Psychological Bulletin, 118*(2), 172–177. https://doi.org/10.1037/0033-2909.118.2.172

Benefo, E. O., Tingler, A., White, M., Cover, J., Torres, L., Broussard, C., & Patra, D. (2022). Ethical, legal, social, and economic (ELSE) implications of artificial intelligence at a global level: A scientometrics approach. *AI Ethics, 1-16*. https://doi.org/10.1007/s43681-021-00124-6

Bennett, C. J., & Raab, C. D. (2020). Revisiting the governance of privacy: Contemporary policy instruments in global perspective. *Regulation & Governance, 14*(3), 447–464. https://doi.org/10.1111/rego.12222

Biber, D. (2004). If you look at …: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25*(3), 371–405. https://doi.org/10.1093/applin/25.3.371

Biros, D. (2020). "the challenges of new information technology on security, privacy and ethics," *journal of the Midwest Association for Information Systems (JMWAIS)*, 2020, 2. *Article, 1*. https://doi.org/10.17705/3jmwa.000057

Blei, D. M., Ng, A. Y., Jordan, M. I., & Lafferty, J. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022. https://doi.org/10.1162/jmlr.2003.3.4-5.993

Bouma, G. (2009). Normalized (pointwise) mutual information in collocation extraction. *Proceedings of GSCL*, 31–40.

Bromberg, D. E., Charbonneau, É., & Smith, A. (2020). Public support for facial recognition via police body-worn cameras: Findings from a list experiment. *Government Information Quarterly, 37*(1), Article 101415. https://doi.org/10.1016/j.giq.2019.101415

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science, 358*(6370), 1530–1534. https://doi.org/10.1126/science.aap8062

Cadwalladr, C., & Graham-Harrison, E. (2018). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian, 17*, 22 Accessed on 27 January 2022 from: http://freestudio21.com/wp-content/uploads/2018/04/50-million-fb-profiles-harvested-by-cambridge-analitica.pdf.

de Camargo Fiorini, P., Seles, B. M. R. P., Jabbour, C. J. C., Mariano, E. B., & de Sousa Jabbour, A. B. L. (2018). Management theory and big data literature: From a review to a research agenda. *International Journal of Information Management, 43*, 112–129. https://doi.org/10.1016/j.ijinfomgt.2018.07.005

Cao, L., Joachims, T., Wang, C., Gaussier, E., Li, J., Ou, Y., & Subrahmanian, V. S. (2014). Behavior informatics: A new perspective. *IEEE Intelligent Systems, 29*(4), 62–80. https://doi.org/10.1109/MIS.2014.60

Cate, F. H. (2008). Government data mining: The need for a legal framework. *Harv. CR-CLL Rev., 43*, 435.

Caudill, E. M., & Murphy, P. E. (2000). Consumer online privacy: Legal and ethical issues. *Journal of Public Policy & Marketing, 19*(1), 7–19. https://doi.org/10.1509/jppm.19.1.7.16951

Chamola, V., Hassija, V., Gupta, V., & Guizani, M. (2020). A comprehensive review of the COVID-19 pandemic and the role of IoT, drones, AI, Blockchain and 5G in managing its impact. *IEEE Access, 1–1*. https://doi.org/10.1109/ACCESS.2020.2992341

Chatterjee, S. (2019). Impact of AI regulation on intention to use robots. *International Journal of Intelligent Unmanned Systems, 8*(2), 97–114. https://doi.org/10.1108/ijius-09-2019-0051

Chatterjee, S. (2020). AI strategy of India: Policy framework, adoption challenges and actions for government. *Transforming Government: People, Process and Policy*. https://doi.org/10.1108/tg-05-2019-0031. ahead-of-print(ahead-of-print).

Chatterjee, S., Khorana, S., & Kizgin, H. (2021). Harnessing the potential of artificial intelligence to Foster Citizens' satisfaction: An empirical study on India. *Government Information Quarterly, 101621*. https://doi.org/10.1016/j.giq.2021.101621

Chatterjee, S., & Sreenivasulu, N. S. (2019). Personal data sharing and legal issues of human rights in the era of artificial intelligence. *International Journal of Electronic Government Research, 15*(3), 21–36. https://doi.org/10.4018/ijegr.2019070102

Chen, Q., Min, C., Zhang, W., Wang, G., Ma, X., & Evans, R. (2020). Unpacking the black box: How to promote citizen engagement through government social media during the COVID-19 crisis. *Computers in Human Behavior, 106380*. https://doi.org/10.1016/j.chb.2020.106380

Chen, Y. N. K., & Wen, C. H. R. (2021). Impacts of attitudes toward government and corporations on public Trust in Artificial Intelligence. *Communication Studies, 72*(1), 115–131. https://doi.org/10.1080/10510974.2020.1807380

Chen, Y.-N. K., & Wen, C.-H. R. (2020). Impacts of attitudes toward government and corporations on public Trust in Artificial Intelligence. *Communication Studies, 1–17*. https://doi.org/10.1080/10510974.2020.1807380

Cinnamon, J. (2017). Social injustice in surveillance capitalism. *Surveillance and Society, 15*(5), 609–625. https://doi.org/10.24908/ss.v15i5.6433

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management, 60*, Article 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383

Cooke-Davies, T. J., & Arzymanow, A. (2003). The maturity of project management in different industries: An investigation into variations between project management models. *International Journal of Project Management, 21*(6), 471–478. https://doi.org/10.1016/S0263-7863(02)00084-4

Dhillon, G., & Torkzadeh, G. (2006). Value-focused assessment of information system security in organizations. *Information Systems Journal, 16*(3), 293–314. https://doi.org/10.1111/j.1365-2575.2006.00219.x

Di Vaio, A., Hassan, R., & Alavoine, C. (2022). Data intelligence and analytics: A bibliometric analysis of human–artificial intelligence in public sector decision-making effectiveness. *Technological Forecasting and Social Change, 174*, Article 121201. https://doi.org/10.1016/j.techfore.2021.121201

Drmota, M., Szpankowski, W., & Viswanathan, K. (2012, July). *Mutual information for a deletion channel* (pp. 2561–2565). IEEE. https://doi.org/10.1109/ISIT.2012.6283980

Duran, N. D., Hall, C., McCarthy, P. M., & McNamara, D. S. (2010). The linguistic correlates of conversational deception: Comparing natural language processing technologies. *Applied Psycholinguistics, 31*(3), 439–462. https://doi.org/10.1017/S0142716410000068

Dwivedi, Y. K., Ismagilova, E., Hughes, D. L., Carlson, J., Filieri, R., Jacobson, J., & Wang, Y. (2020). Setting the future of digital and social media marketing research: Perspectives and research propositions. *International Journal of Information Management, 102168*. https://doi.org/10.1016/j.ijinfomgt.2020.102168

Dwivedi, Y. K., Kapoor, K. K., & Chen, H. (2015). Social media marketing and advertising. *The Marketing Review, 15*(3), 289–309. https://doi.org/10.1362/146934715X14441363377999

Dwivedi, Y. K., Kelly, G., Janssen, M., Rana, N. P., Slade, E. L., & Clement, M. (2018). Social media: The good, the bad, and the ugly. *Information Systems Frontiers, 20*(3), 419–423. https://doi.org/10.1007/s10796-018-9848-5

e Silva, A. M., Rodrigues, Y. R., & Ishii, R. P.. (2020). RIGOR: A new proposal for predicting infant mortality in government health systems using artificial intelligence in Brazil. *Computer, 53*(10), 69–76. https://doi.org/10.1109/mc.2020.2988626

Engin, Z., & Treleaven, P. (2019). Algorithmic government: Automating public services and supporting civil servants in using data science technologies. *The Computer Journal, 62*(3), 448–460. https://doi.org/10.1093/comjnl/bxy082

European Commission. (2020). Spain AI Strategy Report | Knowledge for policy. https://knowledge4policy.ec.europa.eu/ai-watch/spain-ai-strategy-report_en.

Figenschou, T. U. (2020). Social bureaucracy? The integration of social media into government communication. *Communications, 45*(s1), 513–534. https://doi.org/10.1515/commun-2019-2074

Furumura, M., Iwasa, K., Suzuki, Y., Demachi, T., Ishibe, T., & Matsu'ura, R. S.. (2020). Data retrieval system of JMA analog seismograms in the headquarters for earthquake research promotion of the Japanese government. *Seismological Research Letters, 91*(3), 1403–1412. https://doi.org/10.1785/0220190303

Gerard, F., Imbert, C., & Orkin, K. (2020). Social protection response to the COVID-19 crisis: Options for developing countries. *Oxford Review of Economic Policy, 36*(Supplement_1), S281–S296.

Gomez-Marin, A., Paton, J. J., Kampff, A. R., Costa, R. M., & Mainen, Z. F. (2014). Big behavioral data: Psychology, ethology and the foundations of neuroscience. *Nature Neuroscience, 17*(11), 1455–1462. https://doi.org/10.1038/nn.3812

Gonzalez, R. A., Ferro, R. E., & Liberona, D. (2020). Government and governance in intelligent cities, smart transportation study case in Bogotá Colombia. *Ain Shams Engineering Journal, 11*(1), 25–34. https://doi.org/10.1016/j.asej.2019.05.002

Grimmelikhuijsen, S., Jilke, S., Olsen, A. L., & Tummers, L. (2017). Behavioral public administration: Combining insights from public administration and psychology. *Public Administration Review, 77*(1), 45–56. https://doi.org/10.1111/puar.12609

Harari, G. M., Lane, N. D., Wang, R., Crosier, B. S., Campbell, A. T., & Gosling, S. D. (2016). Using smartphones to collect behavioral data in psychological science: Opportunities, practical considerations, and challenges. *Perspectives on Psychological Science, 11*(6), 838–854. https://doi.org/10.1177/1745691616650285

Haug, N., Geyrhofer, L., Londei, A., Dervic, E., Desvars-Larrive, A., Loreto, V., & Klimek, P. (2020). Ranking the effectiveness of worldwide COVID-19 government interventions. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-020-01009-0

Haynes, J., Ramirez, M., Hayajneh, T., & Bhuiyan, M. Z. A. (2017). A framework for preventing the exploitation of IoT smart toys for reconnaissance and exfiltration. In *International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage* (pp. 581–592). Cham: Springer.

Heawood, J. (2018). Pseudo-public political speech: Democratic implications of the Cambridge Analytica scandal. *Information Polity, 23*(4), 429–434. https://doi.org/10.3233/IP-180009

Hiller, J. S., & Bélanger, F. (2001). Privacy strategies for electronic government. *E-government, 200*(2001), 162–198.

Hobolt, S. B., Tilley, J., & Wittrock, J. (2013). Listening to the government: How information shapes responsibility attributions. *Political Behavior, 35*(1), 153–174.

Hollander, A. S., & Icerman, R. C. (1991). Capital budgeting in governments: The feasibility of artificial intelligence technology. *Expert Systems with Applications, 3*(1), 109–116. https://doi.org/10.1016/0957-4174(91)90091-r

Hou, Y., Xiong, D., Jiang, T., Song, L., & Wang, Q. (2019). Social media addiction: Its impact, mediation, and intervention. *Cyberpsychology: Journal of psychosocial research on cyberspace, 13*(1). https://doi.org/10.5817/CP2019-1-4

Hull, R., Kumar, B., Lieuwen, D., Patel-Schneider, P. F., Sahuguet, A., Varadarajan, S., & Vyas, A. (2004). Enabling context-aware and privacy-conscious user data sharing. In *IEEE international conference on Mobile data management, 2004. Proceedings. 2004* (pp. 187–198). IEEE.

Hursh, S. R. (1984). Behavioral economics. *Journal of the Experimental Analysis of Behavior, 42*(3), 435–452.

Irvin, R. A., & Stansbury, J. (2004). Citizen participation in decision making: Is it worth the effort? *Public Administration Review, 64*(1), 55–65. https://doi.org/10.1111/j.1540-6210.2004.00346.x

Isaak, J., & Hanna, M. J. (2018). User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer, 51*(8), 56–59. https://doi.org/10.1109/MC.2018.3191268

Jimenez-Gomez, C. E., Cano-Carrillo, J., & Falcone Lanas, F. (2020). Artificial intelligence in government. *Computer, 53*(10), 23–27. https://doi.org/10.1109/mc.2020.3010043

Jindal, R., & Borah, M. D. (2013). A survey on educational data mining and research trends. *International Journal of Database Management Systems, 5*(3), 53. https://doi.org/10.5121/ijdms.2013.530

Kamolov, S., & Teteryatnikov, K. (2021). Artificial intelligence in public governance. In *Technology and business strategy* (pp. 127–135). Cham: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-63974-7_9.

Kankanhalli, A., Charalabidis, Y., & Mellouli, S. (2019). IoT and AI for smart government: A research agenda. *Government Information Quarterly, 36*(2), 304–309. https://doi.org/10.1016/j.giq.2019.02.003

Kavanaugh, A. L., Fox, E. A., Sheetz, S. D., Yang, S., Li, L. T., Shoemaker, D. J., & Xie, L. (2012). Social media use by government: From the routine to the critical. *Government Information Quarterly, 29*(4), 480–491. https://doi.org/10.1016/j.giq.2012.06.002

Khan, S. M. (2017). Multimodal behavioral analytics in intelligent learning and assessment systems. In *Innovative Assessment of Collaboration* (pp. 173–184). Cham: Springer.

Kraus, S., Breier, M., & Dasí-Rodríguez, S. (2020). The art of crafting a systematic literature review in entrepreneurship research. *The International Entrepreneurship and Management Journal, 1-20*. https://doi.org/10.1007/s11365-020-00635-4

Krippendorff, K. (2013). *Content analysis: An introduction to its methodology* (3rd ed., (3rd ed.,. *2013* pp. 221–250). Thousand Oaks, CA, USA: Sage. https://doi.org/10.2307/2288384

LaBrie, R. C., Steinke, G. H., Li, X., & Cazier, J. A. (2018). Big data analytics sentiment: US-China reaction to data collection by business and government. *Technological Forecasting and Social Change, 130*, 45–55. https://doi.org/10.1016/j.techfore.2017.06.029

Linders, D. (2012). From e-government to we-government: Defining a typology for citizen coproduction in the age of social media. *Government Information Quarterly, 29*(4), 446–454. https://doi.org/10.1016/j.giq.2012.06.003

Litman, L., Robinson, J., & Abberbock, T. (2017). TurkPrime. Com: A versatile crowdsourcing data acquisition platform for the behavioral sciences. *Behavior Research Methods, 49*(2), 433–442. https://doi.org/10.3758/s13428-016-0727-z

Löfgren, K., & Webster, C. W. R. (2020). The value of big data in government: The case of 'smart cities'. *Big Data & Society, 7*(1). https://doi.org/10.1177/2053951720912775

Lukacik, E. R., Bourdage, J. S., & Roulin, N. (2020). Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews. *Human Resource Management Review, 100789*. https://doi.org/10.1016/j.hrmr.2020.100789

MacDougall, C., & Fudge, E. (2001). Planning and recruiting the sample for focus groups and in-depth interviews. *Qualitative Health Research, 11*(1), 117–126. https://doi.org/10.1177/104973201129118975

Macnamara, J. (2015). Creating an "architecture of listening" in organizations. In *The basis of engagement, trust, healthy democracy, social equity, and business sustainability*. Sydney: University of Technology Sydney.

Madnick, S., Johnson, S., & Huang, K. (2019). What countries and companies can do when trade and cybersecurity overlap. *Harvard Business Review, 4*.

Maher, C. S., Hoang, T., & Hindery, A. (2020). Fiscal responses to COVID-19: Evidence from local governments and nonprofits. *Public Administration Review*. https://doi.org/10.1111/puar.13238

Martín, A., & León, C. (2015). Semantic framework for an efficient information retrieval in the E-government repositories. *Advances in Electronic Government, Digital Divide, and Regional Development, 192–213*. https://doi.org/10.4018/978-1-4666-7266-6.ch011

Martin, K. E. (2015). Ethical issues in the big data industry. *MIS Quarterly Executive, 14*, 2.

Mazurek, G., & Małagocka, K. (2019). Perception of privacy and data protection in the context of the development of artificial intelligence. *Journal of Management Analytics, 6*(4), 344–364. https://doi.org/10.1080/23270012.2019.1671243

McEnery, T., & Hardie, A. (2013). The history of corpus linguistics. *Oxford Handbook of the History of Linguistics, 727*, 745.

McHugh, D., Shaw, S., Moore, T. R., Ye, L. Z., Romero-Masters, P., & Halverson, R. (2020). Uncovering themes in personalized learning: Using natural language processing to analyze school interviews. *Journal of Research on Technology in Education, 52*(3), 391–402. https://doi.org/10.1080/15391523.2020.1752337

McKinley, S. K., Fong, Z. V., Udelsman, B., & Rickert, C. G. (2020). Successful virtual interviews: Perspectives from recent surgical fellowship applicants and advice for both applicants and programs. *Annals of Surgery, 272*(3), e192–e196. https://doi.org/10.1097/SLA.0000000000004172

Men, L. R., & Tsai, W. H. S. (2014). Perceptual, attitudinal, and behavioral outcomes of organization–public engagement on corporate social networking sites. *Journal of Public Relations Research, 26*(5), 417–435. https://doi.org/10.1080/1062726X.2014.951047

Mikalef, P., Lemmer, K., Schaefer, C., Ylinen, M., Fjørtoft, S. O., Torvatn, H. Y., & Niehaves, B. (2021). Enabling AI capabilities in government agencies: A study of determinants for European municipalities. *Government Information Quarterly, 101596*. https://doi.org/10.1016/j.giq.2021.101596

Nagtegaal, R. (2021). The impact of using algorithms for managerial decisions on public employees' procedural justice. *Government Information Quarterly, 38*(1), Article 101536. https://doi.org/10.1016/j.giq.2020.101536

Narayanan, A., Huey, J., & Felten, E. W. (2016). A precautionary approach to big data privacy. In *Data protection on the move* (pp. 357–385). Dordrecht: Springer.

Nasseef, O. A., Baabdullah, A. M., Alalwan, A. A., Lal, B., & Dwivedi, Y. K. (2021). Artificial intelligence-based public healthcare systems: G2G knowledge-based exchange to enhance the decision-making process. *Government Information Quarterly, 101618*. https://doi.org/10.1016/j.giq.2021.101618

Natow, R. S. (2020). The use of triangulation in qualitative studies employing elite interviews. *Qualitative Research, 20*(2), 160–173. https://doi.org/10.1177/1468794119830077

Obar, J. A., & Oeldorf-Hirsch, A. (2020). The biggest lie on the internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society, 23*(1), 128–147. https://doi.org/10.1080/1369118X.2018.1486870

Oey, T., Jones, S., Bullard, J. W., & Sant, G. (2020). Machine learning can predict setting behavior and strength evolution of hydrating cement systems. *Journal of the American Ceramic Society, 103*(1), 480–490. https://doi.org/10.1111/jace.16706

Orlikowski, W. J., & Baroudi, J. J. (1991). Studying information technology in organizations: Research approaches and assumptions. *Information Systems Research, 2*(1), 1–28. https://doi.org/10.1287/isre.2.1.1

Palos-Sanchez, P., Saura, J. R., & Martin-Velicia, F. (2019). A study of the effects of programmatic advertising on users' concerns about privacy overtime. *Journal of Business Research, 96*, 61–72. https://doi.org/10.1016/j.jbusres.2018.10.059

Paul, P., & Aithal, P. S. (2020). Informatics: Foundation, nature, types and allied areas—An Educational & Analytical Investigation. *International Journal of Applied Science and Engineering, 8*(1), 01–09.

Pell, B., Williams, D., Phillips, R., Sanders, J., Edwards, A., Choy, E., & Grant, A. (2020). Using visual timelines in telephone interviews: Reflections and lessons learned from the star family study. *International Journal of Qualitative Methods, 19*. https://doi.org/10.1177/1609406920913675

Pencheva, I., Esteve, M., & Mikhaylov, S. J. (2018). Big data and AI – A transformational shift for government: So, what next for research? *Public Policy and Administration, 095207671878053*. https://doi.org/10.1177/0952076718780537

Pencheva, I., Esteve, M., & Mikhaylov, S. J. (2020). Big data and AI–A transformational shift for government: So, what next for research? *Public Policy and Administration, 35*(1), 24–44. https://doi.org/10.1177/0952076718780537

Polat, Z. A., & Alkan, M. (2020). The role of government in land registry and cadastre service in Turkey: Towards a government 3.0 perspective. *Land Use Policy, 92*, Article 104500. https://doi.org/10.1016/j.landusepol.2020.104500

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945–959.

Rayson, P., & Garside, R. (2000, October). Comparing corpora using frequency profiling. In *The workshop on comparing corpora* (pp. 1–6). https://doi.org/10.3115/1117729.1117730

Ribeiro-Navarrete, S., Saura, J. R., & Palacios-Marqués, D. (2021). Towards a new era of mass data collection: Assessing pandemic surveillance technologies to preserve user privacy. *Technological Forecasting and Social Change, 167*, Article 120681. https://doi.org/10.1016/j.techfore.2021.120681

Roberts, L. D. (2015). Ethical issues in conducting qualitative research in online communities. *Qualitative Research in Psychology, 12*(3), 314–325. https://doi.org/10.1080/14780887.2015.1008909

Sarkis, J., Zhu, Q., & Lai, K. H. (2011). An organizational theoretic review of green supply chain management literature. *International Journal of Production Economics, 130*(1), 1–15. https://doi.org/10.1016/j.ijpe.2010.11.010

Saura, J. R., Palacios-Marqués, D., & Iturricha-Fernández, A. (2021). Ethical Design in Social Media: Assessing the main performance measurements of user online behavior modification. *Journal of Business Research, 129*, 271–281. https://doi.org/10.1016/j.jbusres.2021.03.001

Saura, J. R, Ribeiro-Soriano, D., & Iturricha-Fernández, A. (2022). Exploring the challenges of remote work on Twitter users' sentiments: From digital technology development to a post-pandemic era. *Journal of Business Research, 142*, 242–254. https://doi.org/10.1016/j.jbusres.2021.12.052

Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2021a). Setting B2B digital Marketing in Artificial Intelligence-based CRMs: A review and directions for future research. *Industrial Marketing Management, 98*, 161–178. https://doi.org/10.1016/j.indmarman.2021.08.006

Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2021b). Using data mining techniques to explore security issues in smart living environments in twitter. *Computer Communications*. https://doi.org/10.1016/j.comcom.2021.08.021

Saura, J. R., Ribeiro-Soriano, D., & Palacios-Marqués, D. (2021c). Setting privacy "by default" in social IoT: Theorizing the challenges and directions in Big Data Research. *Big Data Research, 25*, 100245. https://doi.org/10.1016/j.bdr.2021.100245

Schreiner, M., Fischer, T., & Riedl, R. (2019). Impact of content characteristics and emotion on behavioral engagement in social media: Literature review and research agenda. *Electronic Commerce Research, 1-17*. https://doi.org/10.1007/s10660-019-09353-8

Shneiderman, B. (2020). Bridging the gap between ethics and practice. *ACM Trans. Interactive Intelligent Syst., 10*(4), 1–31. https://doi.org/10.1145/3419764
Silva, T., Jian, M., & Chen, Y. (2015). Process analytics approach for R&D project selection. *ACM Transactions on Management Information Systems, 5*(4), 1–34. https://doi.org/10.1145/2629436
Silverman, J. (2017). Privacy under surveillance capitalism. *Soc. Res. Int. Q., 84*(1), 147–164.
Skaug Sætra, H. (2020). A shallow defence of a technocracy of artificial intelligence: Examining the political harms of algorithmic governance in the domain of government. *Technology in Society, 62*, Article 101283. https://doi.org/10.1016/j.techsoc.2020.101283
Snelson, C. L. (2016). Qualitative and mixed methods social media research: A review of the literature. *International Journal of Qualitative Methods, 15*(1). https://doi.org/10.1177/1609406915624574, 1609406915624574.
Sørensen, J., & Kosta, S. (2019). Before and after gdpr: The changes in third party presence at public and private european websites. In *The world wide web conference* (pp. 1590–1600).
Stoica, E. A., Pitic, A. G., & Mihăescu, L. (2013). A novel model for E-business and E-government processes on social media. *Procedia Economics and Finance, 6*, 760–769. https://doi.org/10.1016/s2212-5671(13)00200-1
Streletskaya, N. A., Bell, S. D., Kecinski, M., Li, T., Banerjee, S., Palm-Forster, L. H., & Pannell, D. (2020). Agricultural adoption and behavioral economics: Bridging the gap. *Applied Economic Perspectives and Policy, 42*(1), 54–66. https://doi.org/10.1002/aepp.13006
Susanto, H., Yie, L. F., Rosiyadi, D., Basuki, A. I., & Setiana, D. (2021). Data security for connected governments and Organisations: Managing automation and artificial intelligence. In *Web 2.0 and cloud Technologies for Implementing Connected Government* (pp. 229–251). IGI Global. https://doi.org/10.4018/978-1-7998-4570-6.ch011.
Thompson, S., & Warzel, C. (2019, December 19). *Twelve million phones*. Zero Privacy: One Dataset. Retrieved January 16, 2021, from https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html.
Touma, M., Bertino, E., Rivera, B., Verma, D., & Calo, S. (2017). *Framework for behavioral analytics in anomaly identification. In ground/air multisensor interoperability, integration, and networking for persistent ISR VIII* (Vol. 10190, p. 101900H). International Society for Optics and Photonics.
Turner, R. H., & Killian, L. M. (1957). *Collective behavior* (Vol. 3). Englewood Cliffs, NJ: Prentice-Hall.
Van Der Aalst, W. (2016). Data science in action. In *Process mining* (pp. 3–23). Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-662-49851-4_1.
White, C. L., & Boatwright, B. (2020). Social media ethics in the data economy: Issues of social responsibility for using Facebook for public relations. *Public Relations Review, 46*(5), Article 101980. https://doi.org/10.1016/j.pubrev.2020.101980
Wilson, C. (2022). Public engagement and AI: A values analysis of national strategies. *Government Information Quarterly, 39*(1), Article 101652. https://doi.org/10.1016/j.giq.2021.101652
Wong, J. S. (2019). Driverless government: Speculation, citizenship and collective civic intelligence. *Architecture and Culture, 1–17*. https://doi.org/10.1080/20507828.2019.1647960
Wu, M. W., & Su, K. Y. (1993). Corpus-based automatic compound extraction with mutual information and relative frequency count. *Proceedings of Rocling VI Computational Linguistics Conference VI*, 207–216.
Wu, X., Philip, S. Y., Piatetsky-Shapiro, G., Cercone, N., Lin, T. Y., Kotagiri, R., & Wah, B. W. (2003). Data mining: How research meets practical development? *Knowledge and Information Systems, 5*(2), 248–261. https://doi.org/10.1007/s10115-003-0101-1
Xu, T. L., de Barbaro, K., Abney, D. H., & Cox, R. F. (2020). Finding structure in time: Visualizing and analyzing behavioral time series. *Frontiers in Psychology, 11*. https://doi.org/10.3389/fpsyg.2020.01457
Yang, L., Elisa, N., & Eliot, N. (2019). Privacy and security aspects of E-government in smart cities. In *Smart cities cybersecurity and privacy* (pp. 89–102). https://doi.org/10.1016/B978-0-12-815032-0.00007-X
Yang, Q., & Wu, X. (2006). 10 challenging problems in data mining research. *International Journal of Information Technology and Decision Making, 5*(04), 597–604. https://doi.org/10.1142/S0219622006002258
Zato, C., De Luis, A., Bajo, J., De Paz, J. F., & Corchado, J. M. (2011). Dynamic model of distribution and organization of activities in multi-agent systems. *Logic Journal of IGPL, 20*(3), 570–578. https://doi.org/10.1093/jigpal/jzr005
Zeng, S., Hu, Y., Balezentis, T., & Streimikiene, D. (2020). A multi-criteria sustainable supplier selection framework based on neutrosophic fuzzy data and entropy weighting. *Sustainable Development, 28*(5), 1431–1440. https://doi.org/10.1002/sd.2096
Zentall, T. R., Galizio, M., & Critchfield, T. S. (2002). Categorization, concept learning, and behavior analysis: An introduction. *Journal of the Experimental Analysis of Behavior, 78*(3), 237–248. https://doi.org/10.1901/jeab.2002.78-237
Zhang, W., Wang, M., & Zhu, Y. C. (2020). Does government information release really matter in regulating contagion-evolution of negative emotion during public emergencies? From the perspective of cognitive big data analytics. *International Journal of Information Management, 50*, 498–514. https://doi.org/10.1016/j.ijinfomgt.2019.04.001
Zheng, Y., Yu, H., Cui, L., Miao, C., Leung, C., Liu, Y., & Yang, Q. (2020). Addressing the challenges of government service provision with AI. *AI Magazine, 41*(1), 33–43. https://doi.org/10.1609/aimag.v41i1.5195
Zhou, J., Yang, S., Xiao, C., & Chen, F. (2020). Examination of community sentiment dynamics due to covid-19 pandemic: A case study from Australia. *arXiv preprint*. arXiv:2006.12185.
Zhu, L., Chen, P., Dong, D., & Wang, Z. (2021). Can artificial intelligence enable the government to respond more effectively to major public health emergencies? - taking the prevention and control of Covid-19 in China as an example. *Socio-Economic Planning Sciences, 101029*. https://doi.org/10.1016/j.seps.2021.101029
Zhuoxuan, J., Yan, Z., & Xiaoming, L. (2015). Learning behavior analysis and prediction based on MOOC data. *Journal of computer research and development, 52*(3), 614. https://doi.org/10.7544/issn1000-1239.2015.20140491
Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology, 30*(1), 75–89. https://doi.org/10.1057/jit.2015.5
Zuboff, S. (2019a). The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power: Barack Obama's Books of 2019. In E. Profile Brynjolfsson, & A. N. D. R. E. W. Mcafee (Eds.), *The business of artificial intelligence, Harvard Business Review* (pp. 1–20).
Zuboff, S. (2019b). Surveillance capitalism. *Esprit, 5*, 63–77.
Zuiderwijk, A., Chen, Y. C., & Salem, F. (2021). Implications of the use of artificial intelligence in public governance: A systematic literature review and a research agenda. *Government Information Quarterly, 101577*. https://doi.org/10.1016/j.giq.2021.101577

Jose Ramon Saura is Researcher and Associate professor of Digital Marketing in the Business Economics Department at Rey Juan Carlos University, Madrid (Spain). Previously, he held positions and made consultancy at a number of other companies including Google, L'Oréal, Deloitte, Telefónica, or MRM//McCann, among others. He earned an international Ph.D. in Digital Marketing at the Rey Juan Carlos University, while researching at London South Bank University (LSBU) and Harvard University (RCC at Harvard). His research has focused on the theoretical and practical insights of various aspects of User Generated Data and Content (UGD - UGC), with a specific focus around three major research approaches applied to business and marketing: data mining, knowledge discovery, and information sciences. His research has appeared in leading international business, marketing, and information sciences journals.

Domingo Ribeiro-Soriano is a full professor of business administration at the University of Valencia, Spain. He is also the director of the "Entrepreneurship: from bachelor students to entrepreneur Chair", which is sponsored by the Dacsa Group. As a researcher, he has published more than 100 papers in SSCI-ranked journals. Throughout his career, he has edited and contributed to books, authored papers, and delivered keynote speeches at international conferences. He has worked as a guest editor for A and B journals (in WoS by Clarivate Analytics). He has also led several EU-funded projects. Before starting his career in academia, he worked as a consultant at EY (formerly Ernst & Young).

Daniel Palacios-Marqués is Director of the Master in Direction and Management of Digital Businesses in Universitat Politècnica de València. Daniel has published in Journal such as Tourism Management, Annals of Tourism Research, Small Business Economics, Management Decision, International Journal of Technology Management, Cornell Quarterly Management, Services Industries Journal, Service Business, International Entrepreneurship and Management Journal, Journal of Knowledge Management, Journal of Intellectual Capital, International Journal of Sport Policy and Politics, International Journal of Computational Intelligence Systems, International Journal of Innovation Management and International Journal of Contemporary Hospitality Management, Kybernetes, Human Resource Management, International Journal of Project Management, Technological and Economic Development of Economy, Journal of Organizational Change Management. He is currently Associate Editor of Personnel Review Journal. He has been the winner of the 1st Research Prize at the II Convocatòria de Premis de Prospectiva de l'Agència Valenciana d'Avaluació i Acreditació.