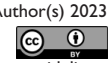
Article

Criminology & Criminal Justice
1–23


Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/17488958231192767
journals.sagepub.com/home/crj

Big Data technologies in criminal investigations: The frames of the members of Judiciary Police in Portugal

Laura Neiva
University of Minho, Portugal

Abstract

Big Data has been increasingly implemented in police departments. In the Judiciary Police in Portugal, although it is at an early stage of implementation, there are no studies on the topic. Based on 16 interviews with members of the Judiciary Police, their frames that portray Big Data's benefits and harms in criminal investigations are explored. The benefits portrayed are related to its capabilities to help fight organized and transnational crimes, advance criminal investigations, and expand the availability of sets of information. The harms focus on the lack of regulatory documents, threats to human rights, and the probability of obtaining erroneous conclusions. A critical analysis of these frames may contribute to reflecting on their role to inform technology developments in policing settings, with implications for inequalities, and crime control.

Keywords

Benefits, Big Data, police officers, risks

Introduction

Since the literature presents various definitions of Big Data (Kitchin, 2014), in criminal investigations, it can be understood as an "amorphous term" that refers to all the digital and technological tools that aim to produce insights and results to help to solve criminal cases. For example, through the analyses and processing of various large datasets (e.g. from criminal databases, and telecommunications metadata) (Brayne, 2018; Ferguson,

Corresponding author:
Laura Neiva, Communication and Society Research Centre (CECS), Institute for Social Sciences, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal.
Email: lauraneiva@ics.uminho.pt

2017; Joh, 2014). Big Data refers to these volume and type of data, and Big Data analytics are the process of exploring and extracting value from these large datasets (Ularu et al., 2012). The majority of Information and Communication Technology (ICT) studies refer to Big Data analytics—a set of techniques commonly conceptualized as link analysis,[1] intelligent agents,[2] text mining,[3] neural network,[4] and machine learning[5] (Oatley, 2022; Pramanik et al., 2017). Given their capabilities in characterizing relationships, identifying (sub)groups, doing real-time control, discovering knowledge from textual data and incomplete datasets, and operating automated processes of statistical operations, they are applied through software such as Analyst's Notebook[6] and InfoGist[7] (Pramanik et al., 2017: 3).

In this article, "Big Data technologies" refers to all these techniques that can be used in the context of "postcrime, when a crime has already been committed" (Simmler et al., 2023: 3) to help find clues to solve criminal cases. In other words, the focus is on all the technologies that enable performing Big Data analytics. This broad conception of "Big Data technologies" is anchored in two main factors. First, in Portugal, in Judiciary Police (PJ), which is the police responsible for conducting criminal investigations in the country (Diário da República, 2000), Big Data technologies are at an early stage of implementation. Thus, specific techniques of Big Data were not known to the participants. Second, as a result of using a wider concept of Big Data, the analysis presented in this article may be applied to other emergent technologies.

Conversely, despite the emergence of this topic in Portugal, considering the European projects ongoing in PJ, it is expected that Big Data will be stated soon. For example, the project "Artificial Intelligence and Advanced Data Analytics for Law Enforcement Agencies" that are focused on cybercrime and terrorism, and it will deliver a descriptive and predictive data analytics platform and related tools which will prevent, identify, analyze, and combat cybercrime and terrorist activities, based on Big Data analytics.[8] Nonetheless, only one study was developed in Portugal about Big Data technologies, still not including police officers' frames. It was a master thesis written by Miguel Pereira (2019) where he did a documentary analysis of the regulatory documents of the Security Police (PSP), stating that all the principles in these documents are related to Big Data. However, although the focus was on policing field, and not in criminal investigations contexts, it was in PSP and not in PJ. Currently (2023), in Portugal, there are no empirical studies about the police frames toward Big Data technologies in criminal investigations.

In the European Union context, we have three studies (Neiva, 2020; Neiva et al., 2022; Simmler et al., 2023) on Big Data technologies in the context of "postcrime" (Simmler et al., 2023: 3). Nevertheless, they are specifically focused on the use of DNA databases (Neiva et al., 2022) in the context of the Prüm System[9] (Neiva, 2020; Neiva et al., 2022) and the Swiss criminal justice system, which includes the use of Big Data not only for criminal investigations purposes, but also to predictive policing, criminal proceedings, forensic-psychiatric assessment, and the penal system (Simmler et al., 2023: 3). In addition, regarding the debate on the benefits and risks of Big Data in criminal investigations, most studies are from the ICT field and are theoretical, not empirical. Nonetheless, they identified some benefits of Big Data in terms of its capabilities to aggregate and analyze different kinds of datasets, and risks, especially on threats to data

protection, bias, and data robustness, associated with the use of Big Data technologies in a criminal investigation (Brady et al., 2015; Iii et al., 2014; Machina and Songjiang, 2020; Pramanik et al., 2017; Ushakov et al., 2021). More specifically, the ICT studies covered the benefits and risks of using Big Data technologies in criminal investigations, but they were not informed by professional frames from police and/or criminal investigations departments.

In contexts wherein Big Data is presented as the future of criminal investigations, further studies on the frames of this technology are required. As noted by Lavorgna and Ugwudike (2021: 2), there is an empirical lack of the narratives that are legitimizing and depicting Big Data technologies, and

> the frames should as such be investigated since they are not simple visualizations of future possibilities: they can inform the design of technologies that either rebalance or maintain an unequal social order and can fuel expansionary applications of the technologies by justice systems.

This article offers an overview of the emergent frames portrayed by members of the PJ in Portugal about the topic by illuminating how these professionals conceive Big Data technologies in terms of potential benefits and harms. Although the benefits and limitations of Big Data are well-documented in scientific literature, "criminological and criminal justice scholarship has paid insufficient attention to the frames and narratives" (Lavorgna and Ugwudike, 2021: 1) that are legitimizing these technologies. This article aims to provide empirical knowledge about police officers' frames—their perceptions—toward Big Data, contributing to research on the narratives about these technologies, and their implications. It may help to reflect on their role to inform technology developments in policing settings, with implications for inequalities, and crime control.

The article's purpose is not to support nor reject the use of Big Data in criminal investigations, much less to sell this technology, nor to convince its implementation under arguments of effectiveness. Instead of that, the goal is to contribute to the Big Data research in criminal investigations literature by: (a) adding knowledge on this topic in an unexplored European country (Portugal), which brings to the academic debate an important asset since most of the discussion is still oriented to the realities of the "global North" and this research is based on a country that occupies a liminal position between the "North" and the "global South"; (b) describing police officers' frames about this technology, which is a topic that appears to be limited explored in the literature; (c) expanding the scientific knowledge about this technology in a "postcrime" (Simmler et al., 2023: 3) context, that is little discussed; and (d) contributing to moving beyond the current scientific debate about specific technologies, with a particular focus on certain Big Data techniques, by adopting as a focus a more generalized conception of Big Data, with the potential to be applied to other types of emerging technologies.

The article is divided into six sections. The first section is the following one, where the current landscape of studies about Big Data technologies to fight crime is highlighted, focusing on empirical studies that include police officers' views about this technology. The second section describes the methodology. The third section presents the frames of members of the PJ in Portugal concerning the use of Big Data in criminal investigations

regarding their narratives about benefits and harms. Finally, the last sections present the discussion, the implications of the study, and the concluding remarks.

Mapping the literature about Big Data technologies and crime

In the last years, with the general societal tendency of datafication—all the information is converted in numbers (Mayer- Schönberger and Cukier, 2013)—that "manifests itself mainly in Big Data analytics" (Lavorgna and Ugwudike, 2021: 1), many criminal and police departments have been doing investments in information technology systems (College of Policing, 2016; Nationale Politie, 2015). Since it is normalized as "a new paradigm in science and society" (Van Dijck, 2014: 198), it has been influencing the criminal justice system (Chan et al., 2022). A report published by Jackson et al. (2014) reveals that police departments are investing in new technological tools, such as Big Data, given their promissory capabilities to improve policing and criminal investigations activities (Chan and Moses, 2017). Although critical criminology and criminal justice scholarship have increasingly developed frameworks to understand the implications of this datafication in crime and social control (see, for example, Lavorgna and Ugwudike, 2021; Smith et al., 2017; Van Dijck, 2014), underlining its consequences on surveillance and control, discrimination and inequality, and privacy and consent, as Prox and Griffiths (2015: 100) stated, "the explosion of 'Big Data' provides a vast new resource that can transform organizations, helping them build smarter systems that drive economic growth, sustainable development, and societal progress" which justifies a growing interest in this technology in the policing field.

The majority of studies about Big Data technologies in crime-fighting activities predominantly discuss the advantages and threats associated with implementing this technology through predictive policing strategies (to predict crimes in terms of location and time; to predict suspects; and to predict victims) (Hardyns and Rummens, 2018; Perry et al., 2013). In terms of police officers' perceptions about the risks and benefits of using this technology to guide predictive police interventions, the studies show the emergence of diverse understandings of Big Data. On one hand, although they conceive these technologies as capable of improving the objectivity and efficiency of their professional activities (Sandhu and Fussey, 2021), providing new techniques to predict crime, do risk assessments, risk management, and crime analysis (Brayne, 2017), they also express contestation in using them (Sandhu and Fussey, 2021). Specifically, given their awareness of the probability of obtaining biased results from a technique that aggregates massive different datasets (Dencik et al., 2018; Sanders et al., 2015; Sandhu and Fussey, 2021). In addition, these reluctant perspectives on Big Data are also emerging from their concerns related to the perception that this technology is not objective since, in their views, if they are controlling systematically the same locals guided by technology insights, they will find more crime or perpetrators in these areas (Sanders et al., 2015; Sandhu and Fussey, 2021).

Therefore, they expressed apprehension about using this technology, given their consciousness about the risks of obtained bias to inform criminal decisions (Dencik et al., 2018; Sanders et al., 2015; Sandhu and Fussey, 2021) and the threats related to the lack

of a regulatory framework focused on Big Data (Fussey and Sandhu, 2022) and privacy and data security (Sanders et al., 2015). To sum up, while police officers conceive the benefits of Big Data in terms of its potential to improve traditional policing strategies (Sanders et al., 2015), they also recognize the risks associated with its implementation (Dencik et al., 2018; Sanders et al., 2015; Sandhu and Fussey, 2021). In addition, Chan et al. (2022: 1–2) demonstrated how datafication affected intelligence production in police services in Australia, New Zealand, Canada, and the United States. Specifically, how "datafication has the potential to transform the epistemic basis of intelligence practice, but such a transformation is likely to be limited and fraught with challenges." Through interviews, they showed that although datafication grew suddenly, traditional methods of policing were simultaneously considered primordial.

However, the majority of studies has as a focus the police officers views' on the capabilities of technologies to enhance the prediction of crime (Babuta, 2017; Brayne, 2017; Dencik et al., 2018; Fussey and Sandhu, 2022; Sanders et al., 2015; Sandhu and Fussey, 2021), instead their views about data technologies uses' to detect the perpetrator of the crimes that have already occurred. In other words, the research about police officers' views on Big Data technologies has been adopted as the focus of their application before a crime is committed (Babuta, 2017; Brayne, 2017; Dencik et al., 2018; Fussey and Sandhu, 2022; Sanders et al., 2015; Sandhu and Fussey, 2021). In the field of criminal investigations, as already stated, only three studies (Neiva, 2020; Neiva et al., 2022; Simmler et al., 2023) addressed the topic. Nonetheless, they explore Big Data technologies from a perspective of "postcrime, when a crime has already been committed" (Simmler et al., 2023: 3) in specific contexts. Namely, Neiva (2020) has written a book on "Big Data na investigação criminal: desafios e expectativas na União Europeia" ["Big Data in criminal investigations: challenges and expectations in the European Union"], where she analyzed the expectations of a set of professionals (professionals working in police cooperation, professionals working on Prüm System, forensic genetics, lawyers, researchers, criminal inspectors, and stakeholders from non-governmental organizations). Her main conclusions stated that, although these professionals conceived Big Data as a powerful tool to advance criminal investigations given its capabilities to analyze and process large amounts of datasets, they also expressed concerns about the threats to human rights, such as privacy and data protection, raised by this technology.

Second, Neiva et al. (2022: 2) analyzed the expectations from professionals working in police cooperation in 16 European Union countries about the potential uses of Big Data in criminal investigations but focused on making use of DNA databases. In this specific context, they show how emerged a "flexible repertoire" of understandings about Big Data that are characterized by perceived benefits—Big Data's potential to solve cold cases and advance the interoperability between different datasets—and perceived risks—difficulties associated with the management of large amounts of information, the genetic discrimination, and threats to privacy and human rights. On the other hand, Simmler et al. (2023), through 34 interviews with criminal justice and police officials, explored the use of algorithms in the Swiss criminal justice system. They present that they are using algorithms for predictive policing, crime analysis, and forensic-psychiatric assessment, given political expectations and demands for efficiency. However, the professionals interviewed reveal perceived benefits that vary between positions of "techno-enthusiastic and techno-skeptics" (Simmler et al., 2023: 19).

Still, these studies are focused on the European context (Neiva, 2020; Neiva et al., 2022) and, in the specific case of Swiss (Simmler et al., 2023), still not in Portugal. Nonetheless, as Neiva et al. (2022) stated, "EU countries have different legislative frameworks and governance principles, technological infrastructures, and organizational imperatives and principles. Such conditions might have an impact on the expectations for Big Data" (Neiva et al., 2022: 10). So, this article aims to study the frames of professionals working in the PJ about Big Data uses in criminal investigations in general, in a specific EU country—Portugal.

Considering the body of research about Big Data technologies in crime-fighting activities, and the proliferation of positive visions around it, there is a need to understand the frames of the professionals that may use this technology in their work practices, and the literature appears to be rather limited. Since the development of this technology and the increasing interest in Big Data, "an actual societal and legal debate on the opportunities and risks must be held" (Simmler et al., 2023: 22). Thus, this article will add further knowledge since it will present a new subject matter about this technology. First, it will address the use of Big Data in criminal investigations after a crime has occurred, and as Neiva et al. (2022) have written, this is an uncharted topic since most studies about Big Data are focused on its predictive capabilities. Second, this study is a part of empirical research with members of the PJ in Portugal, and in this country, until the moment (2023), other studies with these professionals about Big Data technologies do not exist.

Third, as noted by other studies (see, for example, Chan and Moses, 2017), less is known about how police officers perceived Big Data. This article aims to contribute to the knowledge of how these technologies are portrayed by the members of a police force responsible for criminal investigations, taking inspiration from the empirical exploration of frames portraying data-driven technologies for crime prevention and control developed by Lavorgna and Ugwudike (2021: 1). As the authors stated, in the current literature, there are "three frames (optimistic, neutral, oppositional) for understanding how the technologies are portrayed." By using this typology, the analysis presented will explore how police officers are portraying this innovation, and critically reflect on its implications.

Methodology

Empirical data

The empirical data used in this article are a part of a broader research project developed with the PJ in Portugal to understand their expectations about Big Data in criminal investigations. Considering this aim, interviews with the PJ members in Portugal were conducted. PJ is the police force responsible for conducting the country's criminal investigations (Diário da República, 2000).

Sample and recruitment

About the sample, the interviewees were members of PJ in the full and active exercise of their professional activity. This choice was justified by the assumption that they are

social actors who, submerged in a network of meanings characteristic of the professional and social community to which they belong (Albert et al., 2009), play a crucial role in the context under study: they are responsible for criminal investigation. In Portugal, PJ is "clearly dedicated to the criminal investigation" (Durão, 2009: 45), and "it is, per excellence, the entity that manages criminal investigation" (Costa, 2017: 92). Considering it, this sample selection is intentional (Flick, 2002; Starks and Brown Trinidad, 2007; Yin, 2011: 88–89). That is, the participants—members of PJ—were chosen considering that their knowledge and experience enhance and enrich the study of the frames about Big Data technologies in criminal investigations activities.

Regarding the recruitment, the first step was to send requests to the National Director of the PJ to do interviews, by email, by presenting the goals of the research project and, since the research aimed to assess the expectations of using Big Data technologies in criminal investigations, the main request was that the interviewees were in charge of tasks related to criminal investigations. The number of participants was not stated at this stage. The total sample was constructed simultaneously with the development of the research, in a permanent dialogue between the collection and analysis procedures (Guest et al., 2006). For example, initially, after receiving permission from the National Direction to do the interviews, a total of eleven contacts of the PJ members were provided. After that, the second step was to send emails to schedule the interviews. At first, these 11 interviews were done between May and June 2021. Nevertheless, the analysis of these interviews demonstrated the emergence of new topics, and it was necessary to do more interviews. For that, a new request was sent to the National Director of the PJ, and the contacts of five members were provided. These interviews were done in 2022 January, February, and April. The total of interviews was 16. So, the end of the stage of conducting interviews was determined by the saturation point: the moment from which, as more interviews were carried out, the emerging data were neither new nor substantive or, in other words, repeated information was heard during the interviews (Glaser and Strauss, 1967).

Although, as stated earlier, it has not addressed any request about the participants' departments, functions, or hierarchies, they presented a diversified professional characterization. Specifically: one of them was a superior specialist responsible for the statistical analyses of the criminal information; nine criminal inspectors, with responsibilities for conducting operations, diligences, criminal investigations, and procedural acts on ongoing investigations; two criminal investigations coordinators, who are responsible for supervising the criminal investigations activities, planning them and control their legality; one crime analyst with tasks related to the analysis of criminal information to support criminal resources and assist criminal investigations; and, finally, three scientific police specialists, with responsibilities on the managing of the police information system, the availability of criminal information, and the development of police information technology systems through the creation of technological and digital tools, and support their implementation. These different professionals profiles did not trouble the research but, instead of that, were a positive point since, as Morgan (1988: 42) stated, "in organizations, not all interviewees should come from the same hierarchical level or belong to a single department" to access the diversity of the organizational culture.

Interviews

Considering the aim of the study—access how PJ professionals conceive Big Data technologies in terms of potential benefits and harms in criminal investigations—interviews were chosen as the main technique to collect data since they allowed the deepening of questions of meaning and analysis of institutional and social processes, enhancing the study of the phenomenon thoroughly and exhaustively (Starks and Brown Trinidad, 2007). All the interviews were done at the participants' workplaces in PJ. The interviews were semi-structured (Bernard, 2016), with a set of "open" questions that covered the following main themes: (a) Big Data technologies in criminal investigations in Portugal and (b) ethical, legal and social challenges of Big Data in this field. They were, on average, 90 minutes long, conducted in Portuguese, and recorded after all the professionals interviewed signed the authorization and the consent form. This research has an Ethical Protocol to comply with the General Data Protection Regulation 679/2016 to protect the anonymity and confidentiality of the interviewees. To guarantee this was also attributed a number to each interview. This number is used and codifies all the quotations presented. All the interviews were transcribed and translated into English. After that, some editions in the initial transcription were done to delete expressions that were not useful for the analysis and to eliminate parts with the potential to identify the interviewees. This editing process does not mine nor suppress the meanings of the participants' discourses (Poirier et al., 1999).

Analytical strategy

For data analysis, the content analysis in qualitative research (Mayring, 2004) was done to answer the research question: What are the frames portrayed by the members of the PJ about the potential benefits and harms of using Big Data technologies in criminal investigations in Portugal? Specifically, the frames portrayed by the members of the PJ about the potential benefits and harms of using Big Data technologies in criminal investigations in Portugal were first coded after multiple reads to access the diverse perceptions of the interviews. Following the grounded theory, all the excerpts were compared, contrasted, and synthesized in thematic categories (Charmaz, 2006).

To do the analysis, the frames developed by Lavorgna and Ugwudike (2021: 1–3) "for understanding how the technologies are portrayed" were considered. The authors, through a narrative review of abstracts from academic articles on the technologies and their capabilities, advanced the knowledge to understand how the capabilities of data-driven technologies are depicted in the extant literature. The analysis presented in this article was based on their conceptualization of "three frames (optimistic, neutral, oppositional)." Following the authors' approach, data-driven technologies can be portrayed through

> (a) optimistic frames which endorse the tools and their ostensible status as the panacea for cost-effective and efficient crime control; (b) neutral frames that are ambivalent towards the potential harms and benefits of datafication although some are solutionist in that they acknowledge the potential harms of datafication and proffer remedial artificial intelligence (AI) models and (c)

> oppositional frames which emphasize several harms of datafication and reject the view that data-driven tools constitute the panacea for crime control (Lavorgna and Ugwudike, 2021: 1–3).

In the next section, the content analysis is presented. For that, the most illustrative excerpts of the thematic categories were selected.

The neutral frames of the members of PJ in Portugal about Big Data technologies in criminal investigations

In this section, the main findings of the analysis are presented. First, by showing the potential benefits of using Big Data technologies in criminal investigations expressed by the members of the PJ in Portugal. Second, by sharing the potential harms portrayed by the participants on the use of this technology to guide criminal investigations activities. Finally, showing how this ambivalence between benefits and harms demonstrates how they show "neutral frames." These frames are defined as "non-committal in their analysis of the harms or benefits of datafication. These frames evoked a technorealism that vacillated between utopic and dystopic imaginaries and fell into three categories: ambivalent, informational and solutionist" (Lavorgna and Ugwudike, 2021: 8).

Potential benefits

Regarding the potential benefits of using Big Data in criminal investigations, general perceptions are related to the capabilities of this technology to help fight organized and transnational crimes, advance criminal investigations, and expand the availability of sets of information.

The major portrayed benefits are related to the expected potentialities of Big Data to help solve organized and transnational crimes, like terrorism and informatic crimes, that involve a set of countries. As the following inspector interviewed referred:

> [Big Data would be useful for crimes at the level of] terrorism, the informatic crimes, all the crimes that involve more than one state . . . The really big crimes . . . (PJ_2, Inspector)

This first demand of the benefits of Big Data is its potential to help solve organized and transnational crimes, given its capabilities in accessing, process, and correlating different kinds of information. These participants, both inspectors at PJ, indicated that, in their opinions, the big crimes, such as terrorism, sexual crimes, and informatic crimes, require this type of technology with these characteristics (velocity, variety of data, and access to large amounts of information). Given the broad scope, in terms of datasets, that Big Data may access and make available, which can be used by police members, it will be useful to combat transnational crimes.

> Big Data is a lot of information! If we are talking about transnational crimes . . . The trendy crimes like informatic crimes, terrorism, sexual crimes . . . It makes sense this kind of personalized treatment. (PJ_8, Inspector)

A second demand of the potential benefits of Big Data is that it could advance criminal investigations. The following excerpt, from an inspector in PJ in Portugal, highlights how they consider as one of the advantages of Big Data its velocity to process larges amounts of information. It is conceived helpful to cross different kinds of data to support criminal investigations activities.

> I think Big Data technologies have a great utility! Because they allow a faster processing of the volume of information . . . It allows cross and collect information, and sometimes it's very important! It may even help us! (PJ_2, Inspector)

Given that time is an important variable in criminal investigation activities bounded by restricted legal times, the speed of Big Data processing is conceived as a benefit of using this technology in their work. As the following quotation shows:

> Obviously that Big Data will give us an added value in terms of the speed with which this information can eventually be processed because it is also important. Time management is very important in this type of activity [criminal investigation]. (PJ_4, Inspector)

This potential benefit expressed by the inspector interviewed in PJ about the capabilities of Big Data to give their *added value* is according to the perceptions emerging in the study conducted in Swiss, with police officials about their perceptions on algorithms for investigations and crime analysis. As Simmler et al. (2023: 12) stated, "according to all respondents, they [algorithms tools] bring considerable added value in the sense of a qualitative improvement in police work."

Consequently, it is expected, by the interviewees, that Big Data technologies will advance criminal investigations regarding the velocity through making information available:

> [Big Data technologies] would speed up criminal investigation because they could make information available faster. (PJ_15, Scientific Police Specialist)

The last potential benefit expressed by the participants is related to the possibilities created by Big Data technologies, given their capabilities to access a large amount of data and expand the availability of sets of information. The following quotation is illustrative of this emerging perception:

> [Big Data technologies] would increase the amount of information available, with immediate access and eventually enrich the data that the police have! (PJ_15, Scientific Police Specialist)

The potential benefits from the participants in this study are according to the results presented in the studies developed with police officers by Sanders et al. (2015), Babuta (2017), and Sandhu and Fussey (2021) since they argued that their professionals interviewed also expressed positive perceptions about data technologies, in terms of their capabilities to improve the efficiency of policing activities. As the following excerpt demonstrates, this inspector considers that Big Data technologies advance the efficiency of criminal investigations activities since they allow them to be closer to reality:

> I think that criminal investigation is based on obtaining data, so if we have more data, we can be closer to reality. That is why I think that Big Data is very important for criminal investigation! (PJ_5, Inspector)

However, as these potential benefits are not sustained by empirical studies (Meijer and Wessels, 2019), the expression of them by the participants demonstrates how their perceptions about Big Data are framed by the sociotechnical imaginaries about technologies (Lavorgna and Ugwudike, 2021: 3), and grounded by the dominant narrative about datafication (Van Dijck, 2014) which emphasizes their advantages, instead of their threats (Lavorgna and Ugwudike, 2021: 3). If the members of PJ are not using Big Data, and the studies about it are disproportionately focused on their risks instead of their advantages, their potential benefits come from the generalized discourse that circulates in society, with practical implications for Big Data's implementation in police forces, as the power to contribute for their expansion (Lavorgna and Ugwudike, 2021; Smith et al., 2017).

Potential harms

Nonetheless, as stated by Lavorgna and Ugwudike (2021: 8), "most of the neutral frames were nuanced and evoked considerable ambivalence and fluctuation between optimism and pessimism in envisioning social and technoscientific orders." This section of the article presents how the members of the PJ in Portugal also portray some potential harms related to the use of this technology in criminal investigations. Regarding it, the members of the PJ in Portugal answered with topics about their concerns on the use of Big Data in human rights, the lack of regulatory documents to use this technology, and the probability of obtaining erroneous conclusions.

About the potential harms related to the Big Data's threats, the interviewees considered that this technology might impact the human rights of the person who is the subject of the investigation, as stated by the following criminal investigation coordinator interviewed at PJ:

> These platforms such as Big Data can jeopardize the rights, freedoms, and guarantees of the person who is the subject of the investigation. (PJ_10, Criminal Investigation Coordinator)

These potential harms expressed by the participants are according to social sciences literature on the topic as many studies are reporting the threats of Big Data to human rights (Brayne, 2017; McDermott, 2017; Mann and Matzner, 2019; Richardson et al., 2019). Since, in the views of the members of PJ, Big Data deals with private data, one of its risks is exacerbating the challenges of using this kind of personal information:

> Challenges [with Big Data] also exist in the fact that we deal, in criminal investigations activities, with very private data. And therefore that is a legal problem . . . . (PJ_15, Scientific Police Specialist)

These perspectives—on the risks of Big Data to harm human rights—were also presented in the study conducted by Neiva et al. (2022: 8), where they refer that

"interviewees also emphasize the challenges that the uses of Big Data might pose to the protection of sensitive personal data." As the following excerpt illustrates, the participants in this study portray the same potential harm:

> I think that Big Data brings more challenges to data protection and guaranteeing rights! (PJ_3, Superior Specialist)

In addition, the participants in this study also conceived the amount of information produced by Big Data as a benefit and harm. As the following interviews referred,

> I think that Big Data allows us to obtain certain personal and private information . . . It is a huge amount of information processed that can lead to very personal conclusions . . . There are obvious intrusions into privacy . . . Moreover, who has the right to this information? There is only one person who has that right, and that is the person who owns the data. That is why I think Big Data can jeopardize individual rights and individual freedoms . . . . (PJ_5, Inspector)

These perceptions fit the contemporary debate in social sciences literature about the risks associated with Big Data in law enforcement activities since many authors have already pointed out the threats that this technology presents to human rights, specifically to privacy and data protection (Gonçalves, 2017; Herschel and Miori, 2017; Lyon, 2014; Mann and Matzner, 2019; Richardson et al., 2019).

These harms portrayed by the members of the PJ are exacerbated by the current lack of regulatory documents to use this technology. As Pete Fussey (2022: 350) points out, "this growing gap—between digital innovation and meaningful oversight and regulation—is an interstitial space where a lot can happen: biased uses of surveillance, casual attributions of suspicion, unaudited decision making, and unverified biases." The respondents perceived the harms associated with the regulatory absence by emphasizing the need to create a legal document to define how this technology can be used, for whom, and the control of its implementation. These frames align with previous studies realized with police officers (Neiva et al., 2022; Sandhu and Fussey, 2021). Since the existent European legal document, the Directive EU 2016/680, concerning "the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offenses or the execution of criminal penalties" (European Parliament, 2016) that was transposed to Portugal by the Law 59/2019 (2019) does not refer specifically Big Data technologies, it arises risks, and the interviewees consider that these questions should be legally addressed to regulate their use.

> The question is whether legally I can do that [use Big Data technologies] . . . Maybe not . . . I cannot have direct access . . . Now I think that things have to be regulated and within the scope of an investigation, with judicial authorization . . . . (PJ_2, Inspector)

These cautious perceptions about the use of Big Data, given the legal restrictions in terms of access to personal data, underline the results obtained by Simmler et al. (2023: 12) since they described that their interviewees also referred that "there are still legal limits due to data protection requirements."

These perceptions are aligned with the category "solutionist frames" advanced by Lavorgna and Ugwudike (2021: 8). As the authors described, "some neutral frames did not focus solely on developing insights into the limits and potential of datafication in criminal justice. Instead, they evoked a realism that consisted of acknowledging limitations and proffering practical solutions." In this specific case, PJ members are advancing with technical solutions and suggestions to solve the perceived potential obstacles, by sharing solutionist frames, with optimism, as the creation of a regulation for this technology. As is shown by the following quote:

> For the application of Big Data, there should be an institutional protocol. And define who has access to what . . . What criteria allow . . . to be controlled! Who has access to that information? (PJ_5, Inspector)

These "solutionist frames" (Lavorgna and Ugwudike, 2021: 8) emphasize that Big Data technologies can never be used without legal instructions to inspectors about how to access this technology and their information.

> Now there will have to be something here [PJ] regulating Big Data. (PJ_8, Inspector)

Nevertheless, although the members of PJ perceived these potential harms of Big Data, and suggested solutions to solve it, they hold to other fields the responsibility for developing plans, documents, and measures to mitigate them. As other lines of critical inquiry on datafication have shown (see, for example, Iliadis and Russo, 2016; Van Dijck, 2014), usually, these kinds of responsibilities are moved onto other professionals, and the users of data technologies and platforms have "only" to comply with rules defined by others.

Third, another potential harm portrayed by the interviewees is related to their awareness of the probability of obtaining erroneous conclusions when they will use Big Data. Although they consider that the results produced by this technology are *much closer to reality than maybe a criminal investigator would*, they emphasize the possibility of obtaining erroneous insights. As this inspector referred,

> I believe that because of the speed of processing, those probabilities calculated by [Big Data] software are much closer to reality than maybe a criminal investigator would . . . The question is in the use of these final results because one of the things I have also gained from years of experience is that 90% of the time, it is what it looks like, or even more . . . 95% of the time, it is what it looks like . . . the problem is the other 5%! And what I am always afraid of with Big Data is that 5%. That rest . . . this worries me! (PJ_5, Inspector)

This is a category described by Lavorgna and Ugwudike (2021: 8) and presented in neutral frames as ambivalence: "recognizing how certain approaches could also generate 'negative externalities' (. . .)" by underlining "the need for 'cautious optimism.'" The fragility to obtain biased results identified by the interviewees—this ambivalence—is consistent with the literature emerging from empirical studies developed with police officers in other countries regarding technologies. Some research showed how

these professionals recognized that Big Data might calculate erroneous results to inform policing strategies (Dencik et al., 2018; Fussey and Sandhu, 2022; Sanders et al., 2015; Sandhu and Fussey, 2021).

From an interview with an inspector at PJ, the following excerpt reveals how the members of this police force are portraying "informational frames," described as narratives that "highlighted the potential and limits of data-driven technologies, including the capacity of the technologies to produce either positive outputs (e.g. race-neutral risk scores) or negative outputs (e.g. biased risk scores and invasions of privacy)" (Lavorgna and Ugwudike, 2021: 8). By sharing this frame, the following excerpt shows how this interviewee is referring to the probability of obtaining erroneous conclusions with Big Data technologies that may lead to other problems, like the stigmatization and discrimination of individuals through their behaviors. In other words, Big Data's erroneous conclusions may result in discrimination based on algorithms:

> There is an issue here [with Big Data] where we may be running the risk of identifying someone who . . . Just because of their behavior . . . Maybe a potential criminal, but is not until they commit the crime! We cannot judge or detain someone without her first having committed a crime . . . . (PJ_11, Inspector)

Similarly, the quotation presented below also shows how the scientific police specialist portrayed an "informational frame" (Lavorgna and Ugwudike, 2021: 8) by evoking the possibility of obtaining *false positives and false negatives* with the use of Big Data technologies.

> Many of these technologies [like Big Data] are based on statistics, and they never give absolute results . . . They always give probabilities of . . . Very high, always . . . But they are always probabilities, and therefore there are always false positives and false negatives . . . There is always a margin of error . . . . (PJ_15, Scientific Police Specialist)

This potential harm is related to one of the error rates of Big Data identified in other studies (Babuta, 2017; Ferguson, 2016; Herschel and Miori, 2017). As Andrew Ferguson (2016: 1158) defined, false positives occur when Big Data technology produces a result to guide a police intervention to an area or suspect that there did not occur a crime. False negatives are produced when the technology results in the wrong places or persons. Nonetheless, the interviewees do not recognize that this probability to obtain wrong results from the technology is not only a problem that came from this innovation but is also a question raised by the information disproportionality inserted on police databases. As other studies have shown (see, for example, Browning and Arrigo, 2021; Ferguson, 2018), the outputs calculated by Big Data technologies to inform police decisions are based on the information already included in their databases. If this information focuses more on certain social groups, the outputs produced will reflect this over-representation of information. However, the likelihood of getting wrong outputs from using Big Data is perceived by PJ members as a risk that comes from the technology itself and not from the policing system. The misrepresentation of this harm, which is moved by interviewees from policing in general to technology in specific, may impact the perpetuation of

historically biased and discriminatory police practices, exacerbated by the exponential use of technologies.

Finally, this ambivalence between potential harms and benefits also showed how the members of PJ reflect the frames presented in the current literature. As Lavorgna and Ugwudike (2021: 1–8) demonstrated in their narrative review, "the extant literature," although presents other frames, is also "neutral" and "ambivalent toward the potential harms and benefits of datafication," by evoking "a technorealism that vacillated between utopic and dystopic imaginaries," that is shared by the interviewees. These ambivalent perceptions can also emerge from the achievement of the police that the use of Big Data by them may generate mixed public reactions. On one hand, citizens may perceive these technologies as powerful tools for criminal investigations. On the other hand, others may understand Big Data as invasive or prone to misuse. Aware of this, members of PJ present ambivalent arguments to balance the benefits of these technologies with maintaining public trust and addressing concerns about human rights.

Discussion

Most research about Big Data in policing activities is focused on the effectiveness of this technology in predicting crime. While all the studies are very relevant, less is known about how police officers portray this technology. Thus, the aim of this article was not to address the effectiveness of Big Data but rather to explore how future users of this technology perceive its potential harms and benefits in their professional practices of criminal investigation. Accessing these narratives is crucial, as how PJ members understand these technologies is not simply a projection about the future, but has the potential to guide the implementation of Big Data in police departments in Portugal.

The diversity of perceptions emerging from the participants illustrated how the members of the PJ in Portugal are divided between the advantages and the threats of Big Data technologies for criminal investigations purposes. In Lavorgna and Ugwudike's (2021: 2) words, these narratives reflect "neutral frames" that oscillate between the risks and benefits of technologies. These frames "were vacillatory and just as likely to acknowledge sociotechnical imaginaries as they were to refute them" (Lavorgna and Ugwudike, 2021: 11) and are characterized by non-committal in the consideration of advantages and threats of technologies, but fluctuated between the utopian and dystopian of the imaginaries, by not only evoking optimism around technologies but also expressing considerations about caution and some limits of data innovations (e.g. false positives and false negatives), which characterize the potential benefits and harms showed by the participants in this study. These conclusions call for the need to develop, with these social actors, what Lavorgna and Ugwudike (2021: 9) conceptualized as "oppositional frames." Characterized by developing arguments around the "social impacts of datafication," they call attention to "privacy violations, racial bias, algorithmic opacity, fairness concerns, inadequate legislation, the commercialization of datafied criminal justice and systemic bias." As analyzed, the members of PJ also shared some keys from oppositional frames. Nonetheless, they were divided between these aspects and the benefits of the technologies, not only defending an opposition positioning, which reflects their ambition to implement Big Data technologies, although they may recognize their risks. The police

perceptions of threats of Big Data can be understood as a "scapegoat" through which, regardless of their consequences and their risks, these technologies are acclaimed.

This ambivalence is the key point that characterizes his narratives as neutral and not oppositional. In addition, another important point that emerges from this critical analysis of police discourses and underlines the need to develop reflections about the social impacts of datafication with these professionals is the fact that although the risks of Big Data are well-documented by the literature with empirical studies, and the benefits are less empirical known, they also emphasized these benefits. This means that

> the prevalence of optimistic frames we have observed is consistent with current criminal justice deployments of data-driven technologies, which continue to proliferate despite well-documented concerns about data harms. What this indicates is that the frames and their accompanying technologies are central to depictions of the technologies as the panacea for crime control regardless of their accompanying negative impacts (Lavorgna and Ugwudike, 2021: 11–13).

By neutralizing the risks of technologies, and appealing to their benefits, the interviewees demonstrate how "the dominant frames are encouraging the confident expansionary rather than the cautionary datafication of criminal justice" and how they "emphasize the promise of technical accuracy for efficient criminal justice" (Lavorgna and Ugwudike, 2021: 11–13).

These results fit the conclusions from other studies with a focus on perspectives about technologies in policing strategies (Babuta, 2017; Brayne, 2017; Chan and Moses, 2017; Dencik et al., 2018; Fussey and Sandhu, 2022; Sanders et al., 2015; Sandhu and Fussey, 2021; Simmler et al., 2023), criminal investigations (Neiva, 2020; Neiva et al., 2022; Simmler et al., 2023), and intelligence production (Chan et al., 2022). For example, Chan and Moses (2017) also demonstrated how Big Data technologies are conceived as problematic for policing activities but simultaneously helpful. Sandhu and Fussey (2021) showed how police officers' views about technologies are skeptical, but optimist. To sum up, while police officers conceive its risks associated with its implementation (Dencik et al., 2018; Sanders et al., 2015; Sandhu and Fussey, 2021), they also recognize its benefits in terms of improving traditional policing strategies (Sanders et al., 2015).

The participants' potential benefits and harms in this study and the expectations that emerged from other studies including police officers do not present significant differences. These similarities between the frames can be potentially explained by two factors. On one hand, at a micro level, the fact that the interviewees in studies share the same culture, that is, the "police culture". In other words, the community of police officers where they share knowledge, norms, values, and practices, which constitute their professional culture and influence how police officers perform their duties and advocate for their ideas (Cope, 2004; Manning, 1989). Consequently, anchored in the awareness of police to attend to the mixed public reactions on the use of Big Data in criminal investigations.

On the other hand, through the meso and macro levels, the general trend of techno-optimism around the technologies (Quinlan, 2020). Specifically, the widespread optimism around technologies is capable of solving criminal problems, which Andrea

Quinlan (2020: 6) explored in rape kits to sexual crimes and referred to as "techno-optimism, they are generating is tied to a broader faith in technology to rectify the injustices of discriminatory and prejudicial policing, bring justice to victims, and prevent future crime." Furthermore, the expression by PJ members' of the unproven benefits of Big Data shows, not only this techno-optimism, "the influence of positive frames and sociotechnical imaginaries regarding the capacity of technologies to improve human decision-making" (Lavorgna and Ugwudike, 2021: 3), and how their perceptions are grounded by the dominant narrative about datafication (Van Dijck, 2014). These results contribute to the "empirical insight needed to unravel the frames and their implications" (Lavorgna and Ugwudike, 2021: 3), that will be discussed in the next section.

The main limitation of this study is that it does not include interviewees with PSP and Military Police Force (GNR) in Portugal. They are the other police forces also responsible for investigating specific types of crime, but not with the exclusivity of criminal investigation as PJ. Nonetheless, it will be useful to enrich the knowledge about this topic in Portugal if that includes the views from PSP and GNR. Another limitation is related to the relatively small group of interviewees that precludes generalizing, although it is sufficiently adequate for an exploratory study of this nature.

Implications

The conclusions of this study emphasize implications for policy and practice. First, the harms perceived by the participants involved in the study presented in this article are not considering the threats associated with the exponential and uncontrollable growth of technology in police departments with certain interests. In other words, the interviewees obscure other factors related to the adoption of technologies in this field. Namely, political, economic, and capitalist factors that involve the companies responsible to sell this kind of innovation, and are enabled by personal data accessed, shared, and sold without consent (see, for example, Van Dijck, 2014). The threats of Big Data in criminal investigations exceed those perceived by members of PJ, with consequences. For example, although they underlined as a potential benefit of Big Data their capabilities to expand the availability of information, they did not consider the ethics of this data collection, storage, and usage. The neutralization of these harms may have implications for the transparency of the process of access and analysis of the data by the police.

Second, how Big Data technologies are exacerbating the possibilities of datafication and quantification with consequences on reducing a complex phenomenon as is the crime to measurable metrics are also neutralized by the interviewees. This simplification and decontextualization have implications for the understanding of criminality to reduce them. Furthermore, although they mentioned as a risk the errors of Big Data, they did not recognize that this is not only a problem that came from technology but also a question raised by the disproportionality of data in police databases. As other studies have shown (see, for example, Browning and Arrigo, 2021; Ferguson, 2018), the outputs calculated by Big Data technologies to inform police decisions are based on the information already included in their databases. However, the risk of false positives and false negatives with Big Data is perceived by PJ members as a problem that comes from the technology itself and not from the macro policing system. The misrepresentation of this risk, which is

moved by interviewees from policing in general to technology in specific, may have implications to the perpetuation of historically biased and discriminatory police practices, exacerbated by the exponential use of technologies.

Third, even if the interviewees identified as potential harms of Big Data the threats to human rights, the current lack of regulatory documents to use this technology, and the consequences that may have its utilization without an institutional protocol, they demonstrated one tendency underlined by critical inquiry on datafication (see, for example, Iliadis and Russo, 2016; Van Dijck, 2014): they moved the responsibility for developing plans, documents, and measures to mitigate these risks onto other professionals, and positioned themselves as users of data technologies, that have "only" to comply rules defined by others. It reflects that they do not engage critically with these questions.

By contributing to the scholarship on datafication in the criminal justice system, this article emphasizes the need for ethical frameworks and safeguards, by calling for transparency in data collection and analysis processes, and for the establishment of legal and policy frameworks that protects individuals' rights, and emphasizing the need for a critical and reflexive engagement of police officers with datafication, by suggesting how to interrogate its assumptions, consequences, and potential alternatives (Iliadis and Russo, 2016). It can be through the implementation of participative methodologies to discuss these topics (Guston, 2014).

Concluding remarks

In this article, the neutral frames of using Big Data technologies—all techniques that can be used in the context of "postcrime, when a crime has already been committed" (Simmler et al., 2023: 3)—to help find clues to solve criminal cases portrayed by the members of the PJ in Portugal were mapped. The potential benefits described are related to the capabilities of this technology to help fight organized and transnational crimes, advance criminal investigations, and expand the availability of sets of information. At the same time, the potential harms portrayed include concerns with human rights, the lack of regulatory documents to use this technology, and the probability of obtaining erroneous conclusions. The presence of "neutral frames" in PJ members can be understood at the level of micro factors—police culture that is related to negotiations to maintain public trust, addressing, for example, human rights issues; and at the level of macro factors—the techno-optimism that reflects the influence of positive "frames" and sociotechnical imaginaries about the capacities of technologies for crime control in the "frames" of the interviewees. This demonstrates how their perspectives are framed by the dominant narrative of datafication, with the mention of risks being a "scapegoat" because, although recognized, they are neutralized in favor of the proclamation for the implementation of Big Data. These "neutral frames" have implications in terms of their role in influencing the enthusiastic expansion of these technologies in police departments in Portugal, related to the risks that are neutralized and misrepresented by PJ members, with the potential to contribute to inequalities and influence crime control.

This study contributes to the current literature in four ways. First, it has the merit to inaugurate an unexplored research field in Portugal, that is, Big Data technologies in this country's context of criminal investigations; and which brings to the academic debate an

important asset since most of the discussion is still oriented to the realities of "global North" and this research is based on a country that occupies a liminal position between the "North" and the "global South." Second, the evidence described adds further knowledge about the frames of police forces, in terms of benefits and harms, regarding Big Data technologies in criminal investigations, which has received less focus in the literature. Third, given the European projects ongoing in the PJ in Portugal with the development of technologies, it is crucial to understand the professionals' frames that will work with Big Data. Fourth, by adopting a broad conception of "Big Data technologies," the analysis presented in this article may be applied to other emergent technologies.

The aim is not to reject the implementation or development of Big Data in these contexts, but to contribute to an ethical and transparent application of these technologies in criminal investigations. The goal is "that the concerns we raise will heighten awareness among developers and potential users prior to the 'hardening' of the socio-technical structures supporting Big Data" (Moses and Chan, 2014: 645), given that the dissemination of datafication and data technologies as neutral and with benefits by members of PJ has the potential to contribute to "a view of dataveillance as a 'normal' form of social monitoring" (Van Dijck, 2014: 206), since their visions may have power, influence, and a role to inform technology developments in policing settings.

Acknowledgements

The author is deeply grateful to the two anonymous reviewers for their insightful comments that significantly improved the article. Special thanks are also extended to Professor Helena Machado for her invaluable support, encouragement, supervision, and extensive feedback throughout the writing process. Finally, Neiva is thankful to the Judiciary Police in Portugal for authorizing this research, and extends appreciation to all the interviewees for their participation and availability.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work is supported by national funds through FCT—Fundação para a Ciência e a Tecnologia, I.P., under a PhD Research Studentship with the reference 2020.04764.BD (attributed to Laura Neiva) and under the project UIDB/00736/2020 (base funding) and UIDP/00736/2020 (programmatic funding).

ORCID iD

Laura Neiva https://orcid.org/0000-0002-1954-7597

Notes

1. Data mining technique, which reveals the structure of data by representing it as a set of interconnected, linked object or entities (Pramanik et al., 2017: 2).
2. It can perform tasks delegated by their developers, and they are also called task-specific autonomous computing systems (Pramanik et al., 2017: 4).

3. It refers to the process of extracting relevant information and knowledge from textual documents (Pramanik et al., 2017: 5).
4. An artificial neural network is modeled as an aggregation of hundreds of thousands of basic computational units that mimic the function of neurons in a human brain (Pramanik et al., 2017: 5).
5. Used to adapt to dynamic and new circumstances and to detect and extrapolate internal and external structures as well (Pramanik et al., 2017: 6).
6. Assists investigators and analysts by providing faster, more informed decision-making across, and inside organizations. It supplies a centralized, aggregated view of information from different sources (Pramanik et al., 2017: 4).
7. An information retrieval software agent tool that can search and aggregate a set of webpages based on some defined keywords (Pramanik et al., 2017: 5).
8. Source: https://www.policiajudiciaria.pt/projetos-financiados/aida-artificial-intelligence-and-advanced-data-analyitics-for-law-enforcement-agencies/ (last access on 8 June 2022).
9. The Prüm system enables reciprocal automated searching and comparison of DNA profile information (as well as fingerprints and vehicle registration data) for stepping up cross-border cooperation, particularly in combating terrorism and cross-border crime (Machado et al., 2020; Machado and Granja, 2019a, 2019b).

References

Albert M, Laberge S and Hodges B (2009) Boundary-work in the health research field: Biomedical and clinician scientists' perceptions of social science research. *Minerva* 47(2): 171–194.

Babuta A (2017) Big Data and policing an assessment of law enforcement requirements, expectations and priorities. *Royal United Services Institute for Defence and Security Studies*. Available at: https://static.rusi.org/201709_rusi_big_data_and_policing_babuta_web.pdf

Bernard HR (2016) *Research Methods in Anthropology: Qualitative and Quantitative Approaches*. Lanham, MD: Rowman & Littlefield.

Brady O, Overill R and Keppens J (2015) DESO: Addressing volume and variety in large-scale criminal cases. *Digital Investigation* 15: 72–82.

Brayne S (2017) Big Data surveillance: The case of policing. *American Sociological Review* 82(5): 977–1008.

Brayne S (2018) The criminal law and law enforcement implications of Big Data. *Annual Review of Law and Social Science* 14(1): 293–308.

Browning M and Arrigo B (2021) Stop and risk: Policing, data, and the digital age of discrimination. *American Journal of Criminal Justice* 46(2): 298–316.

Chan J and Moses LB (2017) Making sense of Big Data for security. *The British Journal of Criminology* 57(2): 299–319.

Chan J, Sanders C, Bennett Moses L, et al. (2022) Datafication and the practice of intelligence production. *Big Data & Society* 9(1): 1–13.

Charmaz K (2006) *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. Thousand Oaks, CA: SAGE.

College of Policing (2016) National policing vision. Available at: www.college.police.uk/About/Pages/National-policing-vision-2016.aspx (accessed 10 June 2022).

Cope N (2004) Intelligence led policing or policing led intelligence: Integrating volume crime analysis into policing. *The British Journal of Criminology* 44(2): 188–203.

Costa S (2017) O aparato forense e os entendimentos socioculturais na investigação criminal em Portugal. In: Helena M (ed.) *Genética e Cidadania*. Porto: Afrontamento, pp. 87–111.

Dencik L, Hintz A and Carey Z (2018) Prediction, pre-emption and limits to dissent: Social media and Big Data uses for policing protests in the United Kingdom. *New Media & Society* 20(4): 1433–1450.

Diário da República (2000) Decreto-Lei n.O275-A/200, Lei Orgânica da Polícia Judiciária. Available at: https://www.pgdlisboa.pt/leis/lei_mostra_articulado.php?nid=35&tabela=leis&ficha=1&pagina=1 (accessed 25 May 2020).

Durão S (2009) La producción de mapas policiales: prácticas y políticas de la policía urbana en Portugal. *Intersecciones en Antropología* 10(1): 43–61.

European Parliament (2016) Directive 2016/680 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data by competent authorities for the purposes of the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties, and on the free movement of such data, and repealing Council Framework Decision 2008/977/JHA. Available at: http://data.europa.eu/eli/dir/2016/680/oj (accessed 25 May 2020).

Ferguson AG (2016) Policing predictive policing. *Washington University Law Review* 94(5): 1109–1190.

Ferguson AG (2017) *The Rise of Big Data Policing: Surveillance, Race, and the Future of Law Enforcement*. New York: New York University Press.

Ferguson AG (2018) Illuminating Black Data policing. *Ohio State Journal of Criminal Law* 15(2): 503–525.

Flick U (2002) *Métodos qualitativos na investigação científica*. Lisbon: Monitor.

Fussey P (2022) Seeing surveillance: Twenty years of Surveillance & Society. *Surveillance & Society* 20(4): 346–352.

Fussey P and Sandhu A (2022) Surveillance arbitration in the era of digital policing. *Theoretical Criminology* 26: 3–22.

Glaser BG and Strauss AL (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine de Gruyter.

Gonçalves ME (2017) The EU data protection reform and the challenges of Big Data: Remaining uncertainties and ways forward. *Information & Communications Technology Law* 26: 90–115.

Guest G, Bunce A and Johnson L (2006) How many interviews are enough? An experiment with data saturation and variability. *Field Methods* 18(1): 59–82.

Guston DH (2014) Understanding "anticipatory governance." *Social Studies of Science* 44(2): 218–242.

Hardyns W and Rummens A (2018) Predictive policing as a new tool for law enforcement? Recent developments and challenges. *European Journal on Criminal Policy and Research* 24(3): 201–218.

Herschel R and Miori VM (2017) Ethics & Big Data. *Technology in Society* 49: 31–36.

Iii FH, Parker M and Wydra C (2014) The digital case file: The future of fighting crime with Big Data. *Issues in Information Systems* 15: 257–266.

Iliadis A and Russo F (2016) Critical data studies: An introduction. *Big Data & Society* 3(2): 1–7.

Jackson BA, Greenfield VA, Moral AR, et al. (2014) Police department investments in information technology systems. RAND Research report, RAND Corporation, Santa Monica, CA.

Joh EE (2014) Policing by numbers: Big data and the fourth amendment. *Washington Law Review* 89(1): 35–68.

Kitchin R (2014) Big Data, new epistemologies and paradigm shifts. *Big Data & Society* 1(1): 1–12.

Lavorgna A and Ugwudike P (2021) The datafication revolution in criminal justice: An empirical exploration of frames portraying data-driven technologies for crime prevention and control. *Big Data & Society* 8(2): 1–15.

Law 59/2019 (2019) Approves the rules on the processing of personal data for the purposes of preventing, detecting, investigating or prosecuting criminal offenses or enforcing criminal sanctions, transposing Directive (EU) 2016/680 of the European Parliament and of the Council of 27 April 2016. Republic Diary No. 59/2019, Series 1 of 2019-08-08, 41-68. Available at: https://dre.pt/dre/en/detail/act/59-2019-123815983 (accessed 25 May 2020).

Lyon D (2014) Surveillance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data & Society* 1(2): 1–13.

McDermott Y (2017) Conceptualising the right to data protection in an era of Big Data. *Big Data & Society* 4(1): 1–7.

Machado H and Granja R (2019a) Police epistemic culture and boundary work with judicial authorities and forensic scientists: The case of transnational DNA data exchange in the EU. *New Genetics and Society* 38(3): 289–307.

Machado H and Granja R (2019b) Risks and benefits of transnational exchange of forensic DNA data in the EU: The views of professionals operating the Prüm system. *Journal of Forensic and Legal Medicine* 68: 101872.

Machado H, Granja R and Amelung N (2020) Constructing suspicion through forensic DNA databases in the EU. The views of the Prüm professionals. *The British Journal of Criminology* 60(1): 141–159.

Machina AA and Songjiang L (2020) Crime analysis and intelligence system model design using Big Data. *International Journal of Computer Applications* 175: 2212–2221.

Mann M and Matzner T (2019) Challenging algorithmic profiling: The limits of data protection and anti-discrimination in responding to emergent discrimination. *Big Data & Society* 6(2): 1–11.

Manning P (1989) Occupational culture. In: Bailey W (ed.) *Encyclopedia of Police Science*. New York: Routledge, pp. 865–872.

Mayer-Schönberger V and Cukier K (2013) *Big Data*. London: John Murray.

Mayring P (2004) Qualitative content analysis. In: Flick U, von Kardorff E and Steinke I (eds) *A Companion to Qualitative Research*. London: SAGE, pp. 266–269.

Meijer A and Wessels M (2019) Predictive policing: Review of benefits and drawbacks. *International Journal of Public Administration* 42(12): 1031–1039.

Morgan DL (1988) *Focus Groups as Qualitative Research* (Qualitative Research Methods Series). Newbury Park, CA: SAGE.

Moses LB and Chan J (2014) Using Big Data for legal and law enforcement decisions: Testing the new tools. *University of New South Wales Law Journal* 37(2): 643–678.

Nationale Politie (2015) Begroting Nationale Politie 2016-2020. Available at: https://www.parlementairemonitor.nl/9353000/1/j4nvgs5kjg27kof_j9vvij5epmj1ey0/vjxcoseffytb/f=/blg581851.pdf (accessed 25 May 2020).

Neiva L (2020) *Big Data na Investigação Criminal: Desafios e Expectativas na União Europeia* [Big Data in Criminal Investigation: Challenges and Expectations in the European Union]. Famalicão: Húmus.

Neiva L, Granja R and Machado H (2022) Big Data applied to criminal investigations: Expectations of professionals of police cooperation in the European Union. *Policing and Society* 32: 1167–1179.

Oatley GC (2022) Themes in data mining, Big Data, and crime analytics. *WIREs Data Mining and Knowledge Discovery* 12(2): 1–19.

Pereira MR (2019) *Big Data: The case of the strategic information, management and operational control system of the public security police*. MSc Thesis, Higher Institute of Police Sciences and Homeland Security, Lisbon.

Perry WL, McInnis B, Price CC, et al. (2013) Predictive policing: The role of crime forecasting in law enforcement operations. *RAND Corporation*. Available at: https://www.rand.org/content/dam/rand/pubs/research_reports/RR200/RR233/RAND_RR233.pdf (accessed 15 June 2021).

Poirier J, Clapier-Valladon S and Raybaut P (1999) *Life Stories: Theory and Practice*. Oeiras: Celta.

Pramanik MI, Lau RK, Yue WT, et al. (2017) Big Data analytics for security and criminal investigations. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 7(4): 1–19.

Prox R and Griffiths CT (2015) Introduction to the special issue. *Police Practice and Research* 16(2): 99–107.

Quinlan A (2020) The rape kit's promise: Techno-optimism in the fight against the backlog. *Science as Culture* 30(3): 440–464.

Richardson R, Schultz J and Crawford K (2019) Dirty data, bad predictions: How civil rights violations impact police data, predictive policing systems, and justice. *New York University Law Review* 94(2): 192–233.

Sanders C, Christensen T and Weston C (2015) Constructing crime in a database: Big Data and the mangle of social problems work. *Qualitative Sociology Review* 11(2): 180–195.

Sandhu A and Fussey P (2021) The 'uberization of policing'? How police negotiate and operationalise predictive policing technology. *Policing and Society* 31(1): 66–81.

Simmler M, Brunner S, Canova G, et al. (2023) Smart criminal justice: Exploring the use of algorithms in the Swiss criminal justice system. *Artificial Intelligence and Law* 31: 213–237.

Smith GJ, Bennett Moses L and Chan J (2017) The challenges of doing criminology in the Big Data era: Towards a digital and data-driven approach. *The British Journal of Criminology* 57(2): 259–274.

Starks H and Brown Trinidad S (2007) Choose your method: A comparison of phenomenology, discourse analysis, and grounded theory. *Qualitative Health Research* 17(10): 1372–1380.

Ularu EG, Puican FC, Apostu A, et al. (2012) Perspectives on Big Data and Big Data analytics. *Database Systems Journal* 3(4): 3–14.

Ushakov RM, Efremov AD, Gariga OA, et al. (2021) Information technology of Big Data in crime detection, investigation, and prevention. Theoretical and applied analysis of individual prospects and application problems. *Turkish Journal of Computer and Mathematics Education* 12(6): 1645–1657.

Van Dijck J (2014) Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society* 12(2): 197–208.

Yin RK (2011) *Qualitative Research from Start to Finish*, 1st edn. New York: Guilford Press.

Author biography

Laura Neiva is a PhD student in Sociology, under the mentorship of Professor Helena Machado, at the Institute for Social Sciences, Communication and Society Research Centre, University of Minho, benefiting from a PhD Research Studentship supported by the Foundation for Science and Technology (Portugal), with the reference 2020.04764.BD. Neiva's research explores the expectations toward Big Data applied to policing and criminal investigation activities.